

# QUADRATIC-TIME COMPUTATIONS FOR PSEUDO-ANOSOV MAPPING CLASSES

DAN MARGALIT, BALÁZS STRENNER, SAMUEL J. TAYLOR, S. ÖYKÜ YURTTAŞ

ABSTRACT. We give a quadratic-time algorithm to compute the stretch factor and the invariant measured foliations for a pseudo-Anosov element of the mapping class group. As input, the algorithm accepts a word (in any given finite generating set for the mapping class group) representing a pseudo-Anosov mapping class, and the length of the word is our measure of complexity for the input. The output is a train track and an integer matrix where the stretch factor is the largest real eigenvalue and the unstable foliation is given by the corresponding eigenvector. This is the first algorithm to compute stretch factors and measured foliations that is known to terminate in sub-exponential time.

## 1. INTRODUCTION

Let  $S = S_{g,p}$  be the surface obtained from the connected sum of  $g$  tori by deleting  $p$  points. The Nielsen–Thurston classification theorem says that the elements of the mapping class group  $\text{Mod}(S)$  fall into three categories: periodic, reducible, and pseudo-Anosov, and that the last category is exclusive from the other two. To a pseudo-Anosov element  $f$  there are two pieces of data associated: a pair of measured foliations on  $S$  and a real number  $\lambda > 1$ . These are called the stable and unstable foliations of  $f$  and the stretch factor.

The main goal of this paper is to give a quadratic-time algorithm to compute the stretch factor and foliations for a pseudo-Anosov mapping class. The input for the algorithm is a product of generators of  $\text{Mod}(S)$  and the output is a pair  $(\tau, D)$  where  $\tau$  is a train track in  $S$  and  $D$  is a rational matrix acting on the vector space spanned by the branches of  $\tau$ . The stretch factor  $\lambda$  is the largest real eigenvalue of  $D$  and unstable foliation is given by any corresponding positive eigenvector. We refer to the pair  $(\tau, D)$  as a pseudo-Anosov package for  $f$ .

If the word length of the input is  $n$ , we show that the pseudo-Anosov package can be computed using  $O(n^2)$  binary operations. The details of the algorithm are given below, and the quadratic upper bound on the complexity is stated in Corollary 1.2. There is no other algorithm whose running time is known to be faster than exponential. Our paper does not address the issue of computing the eigendata for  $D$ , since there are many well-known (and fast) algorithms for this.

We now explain the broad idea behind the algorithm. The group  $\text{Mod}(S)$  acts in a piecewise-linear fashion on  $\text{MF}(S)$ , the space of measured foliations on  $S$ . Our main technical theorem, Theorem 1.1, says that there is a number  $Q = Q(S)$ , a cell structure on  $\text{MF}(S)$ , and a point  $c$  in  $\text{MF}(S)$  so that for any pseudo-Anosov

---

This material is based upon work supported by the the National Science Foundation under Grant Nos. DMS-1057874, DMS-1811941 DMS-2102018, and DMS-2203431. We also acknowledge support through a Sloan Research Fellowship and a Fulbright Senior Scholar Award.

$f \in \text{Mod}(S)$  the point  $f^Q(c)$  lies in a region of  $\text{MF}(S)$  that has two properties: (1) it contains the unstable foliation for  $f$  and (2) the action of  $f$  on the region is linear. From these two properties we may readily extract the desired pseudo-Anosov package, as the unstable foliation is an eigenvector of the linear map from item (2). We can compute  $f^Q(c)$  and the action of  $f$  at  $f^Q(c)$  in quadratic time; essentially this reduces to the problem of multiplying together  $Qn$  matrices from a given finite list of matrices, where  $n$  is the word length of  $f$ .

Our algorithm has its origins in work of Thurston and Mosher (as explicated by Casson–Bleiler [10, p. 146]), who gave a method for explicitly computing the action of  $\text{Mod}(S)$  on  $\text{MF}(S)$ . In his famous research announcement, Thurston suggested that under iteration by a pseudo-Anosov mapping class, points in  $\text{MF}(S)$  converge to the unstable foliation “rather quickly,” Our Theorem 1.1 is an affirmation of this.

Bell–Schleimer proved a result that, in their words, shows Thurston’s suggestion to be false [4]. Their theorem says that, within the linear region described above, the convergence rate to the ray of unstable foliations is arbitrarily slow. In other words, for the corresponding matrices, the absolute value of the ratio of the two eigenvalues of largest absolute value can be arbitrarily close to 1. There is no contradiction: our result says that the sequence  $c, f(c), f^2(c), \dots$  quickly arrives at the linear region containing the unstable foliation, and the Bell–Schleimer result says that, once there, the convergence of the sequence to the ray of unstable foliations can be slow.

Other works have taken advantage of the fact that the sequence  $f^k(c)$  eventually—and quickly—lands in the “right” linear region for  $f$ ; see the work of Menzel–Parker [20], Moussafir [22], Finn–Thiffeault [13], Hall (who wrote the computer program *Dynn*) [16], the fourth-named author with Hall [17], and Bell (who wrote the computer programs *flipper* and *curver*) [6]. Our Theorem 1.1 is the first result to quantify how quickly the sequence  $f^k(c)$  lands in the right linear region.

There are also algorithms to compute stretch factors and foliations that do not use the action on  $\text{MF}(S)$  directly. Most notably, Bestvina–Handel gave an algorithm that uses train tracks and is based on the Stallings theory of folding [7]. The expectation is that this algorithm has exponential time complexity. Brinkmann [9] implemented the algorithm for once-punctured surfaces in the computer program *Xtrain* and Hall implemented it for punctured disks in the program *Trains*.

We can imagine many possible applications for our algorithm, for instance in determining the smallest stretch factor in  $\text{Mod}(S_g)$  for small  $g$ . Equally important are the various tools we develop in the paper. For example, with our theory of slopes (Section 3) and our forcing lemma (Proposition 5.1) we give a new criterion for a curve to be carried by a train track. We also expect that our work can be used toward a fast algorithm for the conjugacy problem in the mapping class group.

**1.1. The main theorem.** In order to state our main technical result, Theorem 1.1, we require some notation and setup. As above, the surfaces in this paper are all of the form  $S = S_{g,p}$ . Before we start in earnest, we give an abstract topological definition that will be used in the discussion of the space of measured foliations.

*Integral cone complexes.* In this paper, an integral cone complex is a sort of cell complex where a cell is a positive cone in some  $\mathbb{R}^d$  defined by a finite set of integer linear equations and integer inequalities; we will refer to such a cone as an integral linear cone. Such a cell has a finite set of faces, each being a cell in the same sense. Cells are glued along their faces, the gluing maps being linear.

With this basic definition in hand, we have several further definitions to make:

- (a) integral cone complex maps and linear regions,
- (b) the projectivized complex, vertex rays, induced maps, and radius,
- (c) subdivisions for generating sets,
- (d) PL-eigenvectors,
- (e) PL-eigenregions,
- (f) projectively source-sink integral cone complex maps, and
- (g) extended dynamical maps and sink packages.

After making these definitions, we explain how to compute PL-eigenvectors explicitly.

(a) In the category of integral cone complexes, the natural maps  $f : Y \rightarrow Y$  are integral cone complex maps, which we now define. First, suppose  $U$  is a subset of some cell  $V$  of  $Y$ ; we say that a map  $U \rightarrow Y$  is integral linear if the image is contained in a cell  $W$  and—using the  $V$ - and  $W$ -coordinates—the map is given by the restriction of an integral matrix.

With this definition in hand, a map  $f : Y \rightarrow Y$  is an integral cone complex map if there is an integral cone complex subdivision  $Y_f$  of  $Y$  so that the restriction of  $f$  to each cell of  $Y_f$  is an integral linear map. If  $V_i$  is a cell of  $Y$  we denote the cells of such a subdivision of  $V_i$  by  $V_{i,k}^f$ . We will refer to the cells of any such subdivision as linear pieces for  $f$ .

If  $U$  is any integral linear cone contained in a cell  $V$  and  $U$  is also  $f$ -linear, then  $U$  can be extended to a subdivision  $Y_f$  as above. Therefore, it makes sense to refer to any such  $U$  as a linear piece for  $f$ .

We may describe the action of  $f$  on an  $f$ -linear piece  $U$  by a 4-tuple  $(A, V_i, V_j, U)$ , where  $V_i$  and  $V_j$  are cells of  $Y$ , the domain  $U$  is an integral linear cone contained in  $V_i$ , the image  $f(U)$  is contained in  $V_j$ , and  $A$  is the matrix describing the action of  $f$  on the linear piece  $U$  in the given coordinates of  $V_i$  and  $V_j$ .

In our algorithms, it will suffice to record only the triple  $(A, V_i, V_j)$ , omitting the domain. Indeed, by multiplying the matrices from these triples, we can obtain a matrix for the product of mapping classes without knowing the domain. In the case where  $V_i$  is equal to  $V_j$ , we may write  $(A, V_i)$  instead of  $(A, V_i, V_j)$ .

(b) If  $Y$  is an integral cone complex, there is a projectivized version of  $Y$ , denoted  $\mathcal{P}(Y)$ , which is a cell complex in the usual sense. For instance the projectivized complex for  $\text{MF}(S)$  is  $\mathcal{P}(\text{MF}(S)) = \text{PMF}(S)$  (Thurston described the  $\text{Mod}(S)$ -action on  $\text{PMF}(S)$  as opposed to  $\text{MF}(S)$ ). If  $f : Y \rightarrow Y$  is an integral cone complex map, then there is an induced cellular map  $\mathcal{P}(Y) \rightarrow \mathcal{P}(Y)$ . Also, a vertex ray for a integral cone in an integral cone complex  $Y$  is a vertex of the corresponding polyhedron in  $\mathcal{P}(Y)$ .

We now aim to define the radius of  $Y$ , or  $\mathcal{P}(Y)$ , as follows. First, we define a graph  $\mathcal{G}_Y$  whose vertices are the top-dimensional cells  $\mathcal{P}(Y)$  and whose edges connect cells of  $\mathcal{P}(Y)$  that have nonempty intersection. For a set of vertices  $W$  of  $\mathcal{G}_Y$ , we define the radius of  $Y$  at  $W$  to be the supremal distance in  $\mathcal{G}_Y$  from a vertex of  $\mathcal{G}_Y$  to  $W$ . For a vertex  $c$  of  $\mathcal{P}(Y)$ , we define the radius of  $Y$  at  $c$  to be the radius of  $Y$  at  $W_c$ , the set of top-dimensional cells of  $\mathcal{P}(Y)$  containing  $c$ .

As an example, if  $Y$  is any triangulation of  $S^2$  given by the equator and any number of longitudes, then the radius of  $Y$  with respect to the vertex at the north pole is 1. As we will see, there is an analogous situation for  $\text{MF}(S)$ .

(c) Suppose that  $G$  is a group that acts on the integral cone complex  $Y$  by integral cone complex maps, and say that  $G$  is generated by  $\Gamma = \{h_\ell\}$ . We denote by  $Y_\Gamma$  the collection of integral cone complex subdivisions  $\{Y_{h_\ell}\}$  (so there is one subdivision for each generator). As above, we denote the cells of the  $h_\ell$ -subdivision of  $V_i$  by  $V_{i,k}^{h_\ell}$ .

If  $\Gamma$  is finite, then there is a single subdivision (namely, the common refinement of the  $Y_{h_\ell}$ ) with the property that each cell of the subdivision is a linear piece for each  $h_\ell$  and  $h_\ell^{-1}$ . In practice, it is convenient to use a different subdivision for each generator since this reduces the number of calculations required.

(d) Let  $Y$  be an integral cone complex and let  $f : Y \rightarrow Y$  be an integral cone complex map. For a point  $v$  in  $Y$  and a  $\lambda \geq 0$ , the point  $\lambda v$  is a well-defined point of  $Y$ . With this in mind, a PL-eigenvector for  $f$  is a point  $v \in Y$  so that  $f(v) = \lambda v$  for some  $\lambda \geq 0$ . We emphasize that if  $A$  represents  $f$  on a cone  $V$  of  $Y$ , a real eigenvector  $v$  for  $A$  is not necessarily a PL-eigenvector for  $f$ . This is because  $v$  might not lie in the cone of  $\mathbb{R}^d$  corresponding to  $V$ .

(e) Next, a PL-eigenregion for the integral cone complex map  $f$  is a linear piece  $U$  for  $f$  that contains a PL-eigenvector  $v$ . The linear transformation associated to  $f|_U$  will recognize  $v$  as an eigenvector. More specifically, if  $f|_U$  has the triple  $(A, V_i, V_j)$  with  $i = j$  then  $v$  will be a (classical) eigenvector for  $A$ . When  $i \neq j$ , we can choose coordinates on  $V_i$  and  $V_j$  that agree on the cell  $V_{ij} = V_i \cap V_j$ . In these coordinates, we again have that  $v$  is an eigenvector for  $A$ .

As part of our algorithm, we will find a point  $w$  with the property that it is contained in the interior of a union of PL-eigenregions for  $v$ . In such a case, there may be several different triples associated to the action of  $f$  at  $w$ , and hence several different matrices. However, all of these matrices will have  $v$  as an eigenvector.

(f) Next, an integral cone complex map  $f : Y \rightarrow Y$  is projectively source-sink if the induced cellular map  $\mathcal{P}(Y) \rightarrow \mathcal{P}(Y)$  has source-sink dynamics. This means that the induced map has exactly two fixed points, the source and the sink, and that the  $f$ -iterates of all points outside of the source converge to the sink. In particular, the only two PL-eigenvectors for  $f$  (up to scale) are representatives of the source and the sink.

(g) Let  $Y$  be an integral cone complex and  $f : Y \rightarrow Y$  a projectively source-sink integral cone complex map. Let  $U$  be a PL-eigenregion for  $f$  with triple  $(A, V_i, V_j)$ . There exists a maximal subcone  $W$  of  $V_i \cap V_j$  invariant under  $f$ , and the map  $f|_W$  is a dynamical map in the sense that  $f$  maps  $W$  to itself. The cone  $W$  is nontrivial since it contains the ray corresponding to the sink. In Section 1.2 we give an algorithmic procedure to find  $W$  and to compute the linear map associated to  $f|_W$ . We also show that, in any  $\mathbb{R}^d$ -coordinates on  $V_i$ , the cone  $W$  is the intersection of  $V_i$  with a subspace of  $\mathbb{R}^d$ . Hence it makes sense to project  $V_i$  onto  $W$ .

We prove in Lemma 7.3 that (under a certain condition about the location of the sink) the largest real eigenvalue of  $f|_W$  is the eigenvalue for the sink, and

any corresponding positive eigenvector is the sink. If we precompose  $f|W$  with a projection map  $V_i \rightarrow W$  (really a projection of the corresponding subspaces of  $\mathbb{R}^d$ ), we obtain a linear map  $V_i \rightarrow W$ , which we call an extended dynamical map for the sink; say that  $D$  is a matrix for this map. The nonzero real eigenvalues of  $D$  are the same as that of  $f|W$ . As such, the pair  $(V_i, D)$  contains the data of the sink; we refer to any such pair as a sink package.

*The space of measured foliations...* Turning back now to the mapping class group, we denote by  $\text{MF}(S)$  the space of measured foliations on  $S$ , again up to Whitehead moves and isotopy. See the book by Fathi–Laudenbach–Poénaru [12] for background on  $\text{MF}(S)$ .

Let  $S = S_{g,p}$  and let  $\xi(S) = 3g - 3 + p$ . Thurston proved that  $\text{MF}(S)$  is homeomorphic to  $\mathbb{R}^{2\xi(S)}$ . He further proved that  $\text{MF}(S)$  has the structure of an integral cone complex [29]. Thurston actually described the projectivized complex for  $\text{PMF}(S)$ , which can be described in terms of train track coordinates as follows.

A train track in  $S$  is a certain type of embedded finite graph in  $S$  with a well-defined tangent line at each point; see the book of Penner–Harer [24] for details. Thurston’s cells in the cell structure for  $\text{PMF}(S)$ —equivalently, the projectivized integral cone complex structure for  $\text{MF}(S)$ —correspond naturally to train tracks in  $S$ . Conversely, every train track gives a cone in  $\text{MF}(S)$  by considering the measured foliations that are carried by  $\tau$ . Given a train track  $\tau$ , the associated cone will be denoted  $V(\tau)$ . For any given  $S$ , we may choose once and for all a cell decomposition  $Y$  of  $\text{MF}(S)$  with cells  $V(\tau_1), \dots, V(\tau_N)$ .

There are infinitely many choices for the set of  $V(\tau_i)$ , that is, infinitely many choices of structures as an integral cone complex where the cells correspond to train tracks. But they all give the same PL structure. While there are integral cone complex structures on  $\text{MF}(S)$  that do not arise from train tracks, we will only make use of the ones that do; in what follows we refer to these as train track cell structures.

There is a particularly convenient class of train track cell structures on  $\text{MF}(S)$ , namely, the ones arising from Dehn–Thurston train tracks. These are defined as follows. First, we choose a pants decomposition of  $S$ ; we will refer to the curves of the pants decomposition as base curves. From there, we decompose  $S$  into a collection of annuli and pairs of pants. The core curves of the annuli are exactly the base curves and the pairs of pants are the closures of the complementary regions. Each pair of pants has between one and three boundary components, with the other holes accounted for by punctures. As in the book by Penner–Harer [24, §2.6], there are four model (sub)-train tracks on each annulus and at most four on each pair of pants. By gluing together all possible models, we obtain the collection of Dehn–Thurston train tracks. We prove in Lemma 7.2 that if  $Y$  is a Dehn–Thurston cell structure on  $\text{MF}(S)$  and  $c$  is one of the base curves, then the radius of  $Y$  at  $c$  is 1.

*...and the mapping class group action.* There is a natural action of  $\text{Mod}(S)$  on  $\text{MF}(S)$ ; again see [12]. As explained by Thurston, the action of  $\text{Mod}(S)$  on  $\text{MF}(S)$  is by integral cone complex maps (in his setup, piecewise integral projective maps).

Given an integral cone complex structure  $Y$  for  $\text{MF}(S)$  and a finite generating set  $\Gamma$  for  $\text{Mod}(S)$ , there is (as above) an associated collection of subdivisions  $Y_\Gamma$ . Moreover, each subdivision is a train track cell structure. If  $V(\tau_i)$  is a top-dimensional cell of  $Y$  and  $h_\ell$  is an element of  $\Gamma$ , we denote by  $V(\tau_{i,1}^{h_\ell}), \dots, V(\tau_{i,N}^{h_\ell})$  the cells of the

corresponding subdivision of  $V(\tau_i)$ . Each  $\tau_{i,j}^{h_\ell}$  is obtained from the corresponding  $\tau_i$  by a sequence of splittings; again see [24] for details.

An additional property of the  $\text{Mod}(S)$ -action is that each pseudo-Anosov element acts on  $\text{MF}(S)$  with projective source-sink dynamics. In particular, such an element has exactly two PL-eigenvectors. The eigenvalues are  $\lambda$  and  $1/\lambda$ , for the source and sink, respectively. We will refer to a PL-eigenregion for either of these PL-eigenvectors as a Nielsen–Thurston eigenregion.

From a Nielsen–Thurston eigenregion, we obtain an extended dynamical map, and a corresponding sink package  $(V(\tau_i), D)$ , or simply  $(\tau_i, D)$ . Any such sink package  $(\tau_i, D)$  will be called a Nielsen–Thurston package. This is the desired output of our algorithm.

*Statement of the main theorem.* Let  $S = S_{g,p}$ . For the statement of the main theorem, we must define a constant  $Q = Q(S)$ . We build up the definition of  $Q$  by defining a sequence of preliminary constants. We do this in order to give a sense of how  $Q$  arises, and to make feasible to improve our  $Q$  by by improving the intermediate constants (and hence improve the implementation). In the end, we give a definition of  $Q$  using the intermediary constants, and also a formula given directly in terms of  $g$  and  $p$ ; the impatient reader is encouraged to skip to the latter definition of  $Q$ .

We begin with a basic constant related to the topology of  $S$ :

$$K(S) = 4(3g - 3 + p).$$

The number  $K = K(S)$  is the maximum number of separatrices for a measured foliation on  $S$ ; we refer to  $K$  as the separatrix complexity of  $S$ .

Next we have the constants

$$D(S) = 1 \quad \text{and} \quad E(S) = 14.$$

The number  $D = D(S)$  is the radius (as defined above) of any Dehn–Thurston integral cone complex structure on  $\text{MF}(S)$ , specifically the radius about one of the curves of the underlying pants decomposition; see Lemma 7.2. The number  $E(S)$  is an upper bound on the diameter in  $\mathcal{C}(S)$  of the set of vertex cycles for a train track in  $S$ . Aougab [2, Section 5] proved that for all but finitely many  $S$  there is an upper bound of 3 for all train tracks in  $S$ . As observed by Tang–Webb [28, Section 2.3], it follows from the work of Bowditch [8] that there is an upper bound of 14 for all train tracks in all surfaces. Masur–Minsky [19] had previously conjectured that the diameter is bounded above by 3 for all train tracks in all surfaces, but this is still an open problem.

From the above constants, we now define several constants that appear in the statements of the main propositions in the paper or in the proof of Theorem 1.1: Next, we have

$$\begin{aligned} Q_{\text{curve}}(S) &= 2K^2 \\ Q_{\text{track}}(S) &= E \cdot Q_{\text{curve}} \\ Q_{\text{force}}(S) &= Q_{\text{track}} + (2K + 6) \cdot 2K \\ Q_{\text{fit}}(S) &= Q_{\text{force}}(S) + (E + 3) \cdot Q_{\text{curve}} \\ Q_{\text{DT}}(S) &= (D + 1)Q_{\text{track}} \end{aligned}$$

The first three of these constants appear in the statements of Propositions 3.1, 5.1 and 6.1, which we call the bounded slope lemma, the forcing lemma, and the fitting lemma. The number  $2K + 6$  appears in the statement of Lemma 4.1.

Finally, we define

$$Q(S) = Q_{\text{force}}(S) + Q_{\text{fit}}(S) + Q_{\text{DT}}(S) + 1.$$

This gives

$$Q(S_{g,p}) = 2,464\xi^2 + 96\xi + 1$$

and in particular

$$Q(S_2) = 22,465 \quad \text{and} \quad Q(S_3) = 89,281.$$

The coefficient 2,464 arises as  $32(5E+7)$ . If one proves the Masur–Minsky conjecture that  $E(S)$  can be replaced with 3, then we obtain

$$Q(S_{g,p}) = 704\xi^2 + 96\xi + 1, \quad Q(S_2) = 6,625, \quad \text{and} \quad Q(S_3) = 25,921.$$

We may now state our main theorem in terms of Dehn–Thurston cell structures, Nielsen–Thurston eigenregions, and the number  $Q = Q(S)$  just defined.

**Theorem 1.1.** *Fix a surface  $S = S_{g,p}$ , a generating set  $\Gamma$  for  $\text{Mod}(S)$ , a Dehn–Thurston cell structure  $Y$  of  $\text{MF}(S)$ , and a base curve  $c \in \text{MF}(S)$  for the Dehn–Thurston coordinates. Let  $Q = Q(S)$ . Let  $f \in \text{Mod}(S)$  be pseudo-Anosov with unstable foliation  $\mathcal{F}$  and stretch factor  $\lambda$ . Then*

- (1)  $f^Q(c)$  lies in the interior of a union of Nielsen–Thurston eigenregions for  $\mathcal{F}$ , and
- (2) if  $(V(\tau_i), D)$  is a Nielsen–Thurston package for  $f$ , then  $\lambda$  is the largest real eigenvalue of  $D$  and any corresponding positive eigenvector is a positive multiple of  $\mathcal{F}$ .

We may think of the first statement of Theorem 1.1 as saying that the Nielsen–Thurston eigenregion is large in some sense (Proposition 6.1 even more directly has this interpretation). For a general PL-action of a group on a manifold, we would expect the eigenregion for a sink to become arbitrarily small as the word length of  $f$  becomes increasingly large. As such, Theorem 1.1 is not a general fact about PL-actions; our arguments rely heavily on technology specific to the setting of the mapping class group.

The point of the first statement of Theorem 1.1 is that  $f^Q(c)$  finds a Nielsen–Thurston eigenregion for  $f$  (or perhaps several). From the action of  $f$  on any such eigenregion we may compute a Nielsen–Thurston package  $(\tau, D)$ . Then the second statement of the theorem allows us to extract the stretch factor and unstable measured foliation for  $f$ .

We emphasize that the cell structure on  $\text{MF}(S)$  is fixed once and for all. If the cell structure—and hence the notion of linearity—were allowed to depend on  $f$ , then the corresponding statement would hold for  $Q = 1$ , because there exists an invariant train track for  $f$  that carries the curve  $c$ . We also emphasize that our exponent  $Q$  does not depend on the generating set for  $\text{Mod}(S)$ . We will return to this point in Section 1.2.



**1.2. The algorithm.** We now explain in detail the algorithm that comes from Theorem 1.1. As above, the algorithm runs in quadratic time with respect to the length of the spelling of  $f$  as a word in any fixed generating set for  $\text{Mod}(S)$ . And the output is a train track and a matrix. We state this result as Corollary 1.2 below.

There are three precursors to the main algorithm: ledgers, the basic computation algorithm, and the extended dynamical matrix algorithm. We explain these three steps in turn. All three will be utilized in the main algorithm.

*Ledgers: setup for the main computations.* Let  $M$  be a space that admits an integral cone complex structure  $Y$ , and let  $G$  be a finitely generated group acting on  $M$  by integral cone complex maps. In order to perform the algorithm, we make a one-time finite set of choices and computations. The results of these choices and computations can then be used to do computations for any element of  $G$ . First, the choices:

- (1) a generating set  $\Gamma = \{h_1, \dots, h_m\}$  for  $G$  and
- (2) a cell structure for  $Y$  on  $M$  with cells  $\{V_1, \dots, V_N\}$ .

Now the one-time computations:

- (3) the set of subdivisions  $Y_\Gamma = \{Y_{h_\ell}\}$  with the cells of  $Y_{h_\ell}$  denoted  $V_{i,k}^{h_\ell}$  and
- (4) the integer matrices  $\mathcal{A} = \{A_{i,k}^{h_\ell}\}$  so each  $h_\ell|V_{i,k}^{h_\ell}$  is described by  $A_{i,k}^{h_\ell}$ .

We refer to any such quadruple  $(\Gamma, Y, Y_\Gamma, \mathcal{A})$  as a *ledger* for the action of  $G$  on  $M$ . Given such a ledger, we will use the notations  $h_i$ ,  $V_i$ ,  $V_{i,k}^{h_\ell}$ , and  $A_{i,k}^{h_\ell}$  for the elements of  $\Gamma$ ,  $Y$ ,  $Y_\Gamma$ , and  $\mathcal{A}$ , respectively. As above,  $\Gamma$  and  $Y$  are chosen, while  $Y_\Gamma$  and  $\mathcal{A}$  are derived from those choices.

Again, for the case of the  $\text{Mod}(S)$ -action on  $\text{MF}(S)$ , the cells of  $Y_\Gamma$  correspond to train tracks. So each  $V_{i,k}^{h_\ell}$  is equal to some  $V(\tau_{i,k}^{h_\ell})$ , where  $\tau_{i,k}^{h_\ell}$  is obtained from train track  $\tau_i$  by a sequence of splittings; see the book by Penner–Harer for the definition of a splitting [24].

There are many examples of ledgers already in the literature. In his thesis [26], Penner gave a ledger in terms of Dehn–Thurston train track coordinates (explicit matrices are not given, but they can be derived from the computations there). Penner’s formulas contain minor mistakes, corrected in his subsequent paper [23]

*Basic computational algorithm.* Fix a ledger  $(\Gamma, Y, Y_\Gamma, \mathcal{A})$  as above for the action of a finitely generated group  $G$  on a space  $M$  that admits the structure of an  $m$ -dimensional integral cone complex. Suppose we are given some word

$$w = f_n \cdots f_1$$

in the generating set  $\Gamma = \{h_i\}$ , and let  $c \in M$  be a point, written as an integer point in some  $V_i$ . Say that  $w$  represents an element  $f \in G$ . As a precursor to our main algorithm, we would like to be able to compute:

- $f(c) \in M$  and
- a triple  $(A, V_i, V_j)$  describing  $f$  in a partial neighborhood of  $c$ .

By the latter, we mean (as above), a triple  $(A, V_i, V_j)$  where  $V_i$  and  $V_j$  are cells of  $Y$  and  $A$  is a matrix describing  $f$  on a linear piece  $U \subseteq V_i$  containing  $c$  (the matrix  $A$  uses the  $V_i$  and  $V_j$  coordinates). We now give the algorithm for performing these computations.

*Setup.* A space  $M$  and group  $G$  as above, with ledger  $(\Gamma, Y, Y_\Gamma, \mathcal{A})$ .

*Input.* A word  $w = f_n \cdots f_1$  in the generating set  $\Gamma$  (so each  $f_\ell$  is some  $h_{i_\ell}^{\pm 1}$ )



*Initialization.* Let  $c_0 = c$ , let  $A_0 = I_m$  (the identity matrix), and let  $\ell = 1$ .

- (1) *While*  $1 \leq \ell \leq n$ :
  - (a) Determine a top-dimensional cell  $V_{i,k}^{f_\ell}$  of  $Y_\Gamma$  so  $c_{\ell-1} \in V_{i,k}^{f_\ell}$  and  $f_\ell|_{V_{i,k}^{f_\ell}}$  is linear with codomain  $V_j$  and matrix  $A_{i,k}^{f_\ell}$ .
  - (b) Compute  $c_\ell = A_{i,k}^{f_\ell} c_{\ell-1}$  and  $A_\ell = A_{i,k}^{f_\ell} A_{\ell-1}$ .

*Output.*  $f(c) = c_n$  and  $(A_n, V_i, V_j)$ , where  $A = A_n$ , where  $V_i$  contains the  $V_{i,k}^{f_1}$  used in the first step for  $\ell = 1$ , and where  $V_j$  is the  $V_j$  found in the second step for  $\ell = n$ .

We emphasize that there can be different outputs for the triple  $(A, V_i, V_j)$ , since at each stage there may be more than one linear piece of  $f_\ell$  containing  $c$ . We also point out that the creation of the ledger is the step where the generating set for  $G$  is most relevant to the complexity of the algorithm; if we choose a generating set that is not well-suited to the cell structure on  $Y$ , then the entries of the  $A_{i,k}^{h_\ell}$  will be larger.

*The sink package algorithm.* We now give an algorithm to find an extended dynamical matrix and sink package for the sink of a projectively source-sink integral cone complex map  $f : Y \rightarrow Y$ . We suppose that we have in hand a triple  $(A, V_i, V_j)$  for the action of  $f$  on a PL-eigenregion for the sink  $v \in Y$ . Both  $V_i$  and  $V_j$  are cones in  $\mathbb{R}^d$ , and we may assume without loss of generality that whatever faces are shared in  $Y$  are also shared in  $\mathbb{R}^d$ . In this case we say that the coordinates on  $V_i$  and  $V_j$  are compatible.

*Input.* A triple  $(A, V_i, V_j)$  for a PL-eigenregion for the sink of a projectively source-sink integral cone complex map. We assume that the matrix  $A$  is written in terms of compatible  $\mathbb{R}^d$ -coordinates on  $V_i$  and  $V_j$ .

*Initialization.* Let  $W_0$  be the subspace of  $\mathbb{R}^d$  spanned by  $V_i \cap V_j$  (in the compatible coordinates).

- (1) *While*  $1 \leq k \leq m$ :
  - let  $W_k = A(W_{k-1}) \cap W_{k-1}$
- (2) Choose a  $d \times d$  matrix  $B$  that projects  $\mathbb{R}^d$  to  $W$ ; set  $D = AB$ .

*Output.*  $(V_i, D)$

In general, projection onto an integral linear subspace gives a matrix with rational entries. So, generically, we expect the output matrix  $D$  to have rational entries.

The complexity of the sink package algorithm is quadratic in the bit length of the entries of  $A$ . Indeed, the Zassenhaus algorithm converts the problem of computing a basis for an intersection into a row reduction problem, and the complexity of row reduction for integer matrices is known to be quadratic in the bit lengths of the entries; see, e.g. [31, Lecture X, Theorem 5].

We will of course apply the extended dynamical matrix algorithm in the case of the mapping class group, in which case  $V_i$  can be replaced by the train track defining the cell of  $\text{MF}(S)$  corresponding to  $V_i$ . Again, we prove in Lemma 7.3 that the largest real eigenvalue of  $D$  is the eigenvalue for the sink, and any corresponding positive eigenvector is the sink.

*Property Q.* We will state the main algorithm for any group action on an integral cone complex that satisfies the same property as the one guaranteed for the mapping class group by the first statement of Theorem 1.1. Specifically, we say that the action of a finitely generated group  $G$  on an integral cone complex  $Y$  satisfies property Q if there is a natural number  $Q = Q(Y)$  so that, for any  $f \in G$  acting with projectively source-sink dynamics and any  $c \in Y$  lying on a vertex ray, the point  $f^Q(c)$  lies in the interior of a union of PL-eigenregions for  $f$  containing the sink.

*The main algorithm.* Suppose that  $G$  acts by integral cone complex maps on a space  $M$  that admits the structure of an integral cone complex  $Y$ , and let  $f$  be an element of  $G$  that acts on  $M$  with projective source-sink dynamics. Fix a ledger  $(\Gamma, Y, Y_\Gamma, \mathcal{A})$ , and fix some nonzero  $c \in M$  lying on a vertex ray of  $Y$ . Suppose that the  $G$ -action on  $M$  satisfies property Q for some specific  $Q = Q(Y)$ .

*Input.* A word  $w = f_n \cdots f_1$  in the generating set  $\Gamma$  that represents a projectively source-sink element  $f$  of  $G$ .

*Step 1.* Applying the basic computation algorithm to  $w$ , compute  $f^Q(c)$ .

*Step 2.* Applying the basic computation algorithm to  $w$ , compute a triple  $(A, \tau_i, \tau_j)$  associated to an  $f$ -linear piece  $U$  of  $Y$  containing  $f^Q(c)$ .

*Step 3.* Compute a sink package  $(\tau_i, D)$  for  $f$  using  $V(\tau_i)$ -coordinates.

*Output.*  $(\tau_i, D)$

We are finally ready to state our corollary of Theorem 1.1.

**Corollary 1.2.** *Fix a surface  $S$  and let  $\Gamma$  be any finite generating set for  $\text{Mod}(S)$ . The main algorithm computes a Nielsen–Thurston package for any pseudo-Anosov  $f \in \text{Mod}(S)$ , given as a word  $w$  in  $\Gamma$ . The running time of the algorithm is quadratic in the length of  $w$ .*

Corollary 1.2 follows from Theorem 1.1 because if  $w$  has length  $n$ , then Steps 1 and 2 of the main algorithm requires us to multiply  $Qn$  matrices from a given finite list of  $m \times m$  matrices. The number of steps required to perform this multiplication is quadratic in  $Qn$ , hence quadratic in  $n$  (since  $Q$  does not depend on  $w$ ). As above, Step 3 also runs in quadratic time.

We may also be interested in the complexity of the main algorithm with respect to the complexity of the surface,  $\xi(S_{g,p}) = 3g - 3 + p = K(S_{g,p})/4$ . Given a particular surface, we may compute the ledger in exponential time with respect to  $\xi(S)$ , since there are exponentially many cells of the Dehn–Thurston cell structure. Once we have the ledger, we can compute the Nielsen–Thurston package for  $f$  in quadratic time with respect to  $\xi(S)$ , since  $Q$  is quadratic in  $K$ , hence  $\xi$ . As suggested to us by Webb [30], it may be possible to improve our algorithm so that it is also polynomial in  $\xi(S)$ ; one potential approach is to use Penner’s ledger described above.

*The Nielsen–Thurston problem.* The more general problem that our work fits into is the Nielsen–Thurston problem. This is the algorithmic problem of determining, from a product of generators of  $\text{Mod}(S)$ , the Nielsen–Thurston type of the resulting  $f \in \text{Mod}(S)$  and the associated data: the period if  $f$  is periodic, the canonical reduction system if  $f$  is reducible, and the stretch factor and foliations when  $f$  is pseudo-Anosov.

Bell–Webb gave a polynomial-time algorithm for the reducibility problem, that is, the problem of determining whether an infinite order mapping class (given as a word in the generators) is reducible [5]. In this case, their algorithm gives the canonical reduction system. Otherwise, their algorithm computes the asymptotic translation length in the curve graph. (It is known classically that if  $f$  is periodic, then this—plus the period—can be determined in quadratic time. Indeed, this follows from the upper bound on the period (depending only on  $S$ ) and the quadratic-time solution to the word problem for  $\text{Mod}(S)$ .) Thus, combining our work with theirs, we obtain a polynomial-time algorithm to solve the Nielsen–Thurston problem.

Baroni [3] recently gave an alternate version of the Bell–Webb algorithm. Instead of just being polynomial in the word length, his algorithm is polynomial in the word length and the complexity of the surface. On the other hand, Baroni’s algorithm does not explicitly determine any reducing curves.

As pointed out to us by Agol, there is an approach to the Nielsen–Thurston problem using the classical theory of 3-manifolds. Given a mapping class, the algorithm proceeds by producing a triangulation of the corresponding mapping torus, and then using the Jaco–Tollefson algorithm to find the JSJ-decomposition [18, Section 8]. (The key point is that a pseudo-Anosov mapping class has a trivial JSJ-decomposition.)

**1.3. New tools and an outline of the proof.** We now give an outline of the paper and the proofs of the main results, as well as descriptions of the new tools introduced in the paper. We begin in Section 2 by illustrating our algorithm through a concrete example.

The rest of the paper is devoted to the proof of Theorem 1.1. The keys to our proof are Propositions 5.1 and 6.1, called the forcing lemma and the fitting lemma. Both propositions are stated in terms of the notion of slope for a curve in a surface, which is also new. We now turn to developing this idea, beginning with the required background.

*Horizontal and vertical foliations.* Let  $f \in \text{Mod}(S)$  be a pseudo-Anosov mapping class. In order to orient ourselves (and the reader), we will refer to the stable and unstable measured foliations,  $(\mathcal{F}^s, \mu^s)$  and  $(\mathcal{F}^u, \mu^u)$ , as  $(\mathcal{F}_v, \mu_v)$  and  $(\mathcal{F}_h, \mu_h)$ . The ‘ $v$ ’ and ‘ $h$ ’ stand for “vertical” and “horizontal” and in our pictures we will (away from singularities) draw the foliations so that  $\mathcal{F}_v$  is vertical,  $\mathcal{F}_h$  is horizontal, and the transverse measures are  $|dx|$  and  $|dy|$ . In such a picture, the pseudo-Anosov representative of  $f$  stretches horizontally and compresses vertically.

*Singular Euclidean structures.* The two measured foliations  $\mathcal{F}_v$  and  $\mathcal{F}_h$  induce a singular Euclidean structure  $X$  on  $S$ , the singular points being the singularities of the foliations. Let  $\bar{X}$  be the closed singular Euclidean surface defined by taking the completion of  $X$  with the induced metric. There is a set of marked points  $\Sigma \subseteq \bar{X}$  so that  $\bar{X} \setminus \Sigma$  is equal to the surface obtained from  $X$  by deleting the singularities of  $\mathcal{F}_v$  and  $\mathcal{F}_h$ . In other words,  $\Sigma$  is the set of points that either correspond to punctures of  $X$  or singularities of the foliations (or both).

We remark that  $\mathcal{F}_h$  and  $\mathcal{F}_v$  are also the horizontal and vertical measured foliations associated to an integrable meromorphic quadratic differential  $q$  on  $\bar{X}$ , regarded as a closed Riemann surface.

*Section 3: Saddle connections and slope.* Let  $X$  be a singular Euclidean surface. A saddle connection in the surface  $\bar{X}$  is a closed geodesic arc that meets  $\Sigma$  in exactly two points, its endpoints. The corresponding arc in  $X$  is also called a saddle connection.

Let  $\sigma$  be such a saddle connection in  $X$ . We define the *slope* of  $\sigma$  as its height over width:

$$\text{slope}_X(\sigma) = \frac{\mu_h(\sigma)}{\mu_v(\sigma)}.$$

For any collection of saddle connections  $Z$ , we define  $\text{slope}_X(Z)$  to be the set of slopes of elements of  $Z$ .

Let  $c$  be a simple closed curve in  $X$ . By pulling  $c$  tight in  $X$ , the curve  $c$  degenerates to an immersed path in  $\bar{X}$  contained in a set of pairwise non-crossing saddle connections. Here and throughout when we say that saddle connections are non-crossing, we mean that their interiors are disjoint. Similarly, if  $c$  is a measured foliation on  $S$ , we can pull tight each nonsingular leaf. Each leaf gives an immersed path contained in a set of pairwise non-crossing saddle connections. Further, since the leaves are pairwise disjoint, there is a single collection of pairwise non-crossing saddle connections so that each leaf gives an immersed path contained in the union. For  $c$  either a simple closed curve or a foliation, let  $Z = \{\sigma_1, \dots, \sigma_k\}$  be the resulting collection of pairwise non-crossing saddle connections. We define the slope of  $c$  as the finite set

$$\text{slope}_X(c) = \text{slope}_X(Z) = \{\text{slope}_X(\sigma_1), \dots, \text{slope}_X(\sigma_k)\}.$$

Finally, for a train track  $\tau$ , we define  $\text{slope}_X(\tau)$  to be the union of the slopes of the vertex cycles.

*Section 3 (cont.): Horizontality.* It is sometimes convenient for us to use an alternate formulation of the slope. Let  $f : S \rightarrow S$  be a pseudo-Anosov mapping class with stretch factor  $\lambda$  and singular Euclidean surface  $X$ , and let  $c$  (or  $\tau$ ) be as above. We define

$$\mathcal{H}_f(c) = -\frac{1}{2} \log_\lambda \text{slope}_X(c).$$

We refer to  $\mathcal{H}_f(c)$  as the horizontality of  $c$ . We are using an abuse of notation here, as  $\mathcal{H}_f$  is not completely determined by  $f$ . Indeed, different choices of  $X$  lead to horizontality functions that differ by a constant. For any such choice, horizontality satisfies the equation

$$\mathcal{H}_f(f^k(c)) = \mathcal{H}_f(c) + k$$

for any  $k$ . We also prove in Proposition 3.1 that for any curve, arc, or leaf of a foliation  $c$ , the set  $\mathcal{H}_f(c)$  has diameter bounded above by  $Q_{\text{curve}} = 2K^2$ , the key point being that this bound depends only on  $S$  (and not on  $f$ ). One other fact that we will use is that slope and horizontality are continuous functions over  $\text{MF}(S)$ .

For two sets of real numbers  $A$  and  $B$  we say that  $A < B$  if  $\max A < \min B$ . We will use this notion to compare the slopes and horizontalities of the various objects discussed above.

As an application of our upper bound on  $\text{diam } \mathcal{H}_f(c)$ , we give a new approach to, and—in the case of a closed surface—a slight improvement of the result of Gadre–Tsai on stable translation distances in the curve graph; see Section 3.4.

*Section 4: Approximating rectangles.* In Section 4 we prove Lemma 4.1, a technical lemma that will be used to prove the forcing lemma in Section 5. The basic idea of the lemma is to approximate a configuration  $(\gamma_1, \gamma_2, \ell)$ , where the  $\gamma_i$  are crossing saddle connections and  $\ell$  is a horizontal leaf, by a configuration  $(\gamma_1, \gamma_2, R)$ , where  $R$  is a maximal rectangle. For the latter to be an approximation, we mean that the  $\gamma_i$  cross the horizontal sides of  $R$  and that the slope of  $R$ —defined to be the slope of either diagonal—is not too small.

*Section 5: The forcing lemma.* Let  $f$ ,  $\mathcal{F}_h$ , and  $X$  be as above, and let  $\tau$  be a filling, transversely recurrent train track in  $S$ . We write  $c < \tau$  and  $\mathcal{F} < \tau$  to mean that a curve  $c$  or a foliation  $\mathcal{F}$  is carried by  $\tau$ , and in this case we write  $\tau[c]$  and  $\tau[\mathcal{F}]$  for the sub-train track of  $\tau$  consisting of all branches traversed by  $c$  or  $\mathcal{F}$ , respectively.

Suppose that  $c$  is a simple closed curve or a leaf of a measured foliation in  $S$ . In the forcing lemma (Proposition 5.1) we assume that

$$\mathcal{H}_f(c) \geq \mathcal{H}_f(\tau) + Q_{\text{force}}.$$

Under this assumption, the forcing lemma gives the following implications:

- (1) If  $c < \tau$ , then  $\mathcal{F}_h < \tau[c]$ .
- (2) If  $\mathcal{F}_h < \tau$ , then  $c$  is carried by a diagonal extension of  $\tau[\mathcal{F}_h]$ .

Again, for the relevant train track-related definitions, including carrying and diagonal extensions, see the book by Penner–Harer [24].

*Section 6: The fitting lemma.* Suppose  $f$  is a pseudo-Anosov mapping class and that  $\tau_1$  and  $\tau_2$  are two train tracks carrying the horizontal foliation  $\mathcal{F}_h$ . The fitting lemma (Proposition 6.1) gives an  $f$ -invariant train track  $\nu$ —meaning that  $f(\nu) < \nu$ —with two properties:

- (1)  $\nu < \tau_i$  for  $i \in \{1, 2\}$  and
- (2)  $\mathcal{H}_f(\nu) \leq \mathcal{H}_f(\tau_i) + Q_{\text{fit}}$  for  $i \in \{1, 2\}$ .

The train tracks  $\tau_1$  and  $\tau_2$  will be Dehn–Thurston train tracks coming from the cell structure on  $\text{MF}(S)$ . On an intuitive level, if a train track  $\nu$  is carried by  $\tau_i$  and has much larger horizontality, then  $V(\nu)$  is a very small sub-region of  $V(\tau_i)$ . So the  $\nu$  in the conclusion of Proposition 6.1 is fitted in the sense that  $V(\nu)$  is both contained in  $V(\tau_i)$  and is large (intuitively, we think of  $V(\nu)$  as being almost the size of  $V(\tau_i)$ , although it may be of smaller dimension).



FIGURE 1. In the proof of Theorem 1.1, we choose  $Q$  so  $\mathcal{H}_f(f^Q(c))$  is very large, we find a coordinate train track  $\tau_i$  carrying,  $\mathcal{F}_h$ , and we choose an  $f$ -invariant  $\nu$  carried by  $\tau_i$  with  $\mathcal{H}_f(\nu)$  small

*Section 7: proof of the main theorem.* We sketch the idea of the proof of Theorem 1.1. Choose any curve  $c$  lying on a vertex ray of  $\text{MF}(S)$ . Our  $Q$  and  $Q_{\text{DT}}$  are defined exactly so that

$$\mathcal{H}_f(f^Q(c)) \geq \mathcal{H}_f(\tau_i) + Q_{\text{force}} + Q_{\text{fit}} + 1$$

for all train tracks  $\tau_i$  arising in the chosen cell decomposition of  $\text{MF}(S)$ . Suppose that  $f^Q(c)$  is carried by the coordinate train track  $\tau_i$ . By the forcing lemma,  $\mathcal{F}_h$  is carried by  $\tau_i$ . By the fitting lemma, there is an  $f$ -invariant train track  $\nu$  so that

$$\mathcal{H}_f(f^Q(c)) \geq \mathcal{H}_f(\nu) + Q_{\text{force}} + 1$$

Then for every curve  $d$  near  $f^Q(c)$ , the forcing lemma (plus the continuity of horizontality) gives a diagonal extension  $\nu_d$  of  $\nu$  that carries  $d$ . Each such  $V(\nu_d)$  is a Nielsen–Thurston eigenregion. Since the  $V(\nu_d)$  form an open neighborhood of  $f^Q(c)$  in  $\text{MF}(S)$ , every linear piece for  $f$  that contains  $f^Q(c)$  must overlap one of the Nielsen–Thurston eigenregions  $V(\nu_d)$ , and hence is contained in a Nielsen–Thurston eigenregion, as desired.

*Acknowledgments.* We are grateful to Ian Agol, James Belk, Mark Bell, Joan Birman, Tara Brendle, Martin Bridson, Alex Eskin, Benson Farb, Juan González-Meneses, Toby Hall, Chris Leininger, Livio Liechti, Scott MacLachlan, Kasra Rafi, Saul Schleimer, Richard Webb, and Alex Wright for helpful conversations.

## 2. AN EXAMPLE

In this section we apply (a version of) our algorithm to a specific example of a pseudo-Anosov mapping class. Our example lies in the braid group  $B_3$ . This group is naturally isomorphic to the mapping class group of  $D_3$ , the closed disk with three marked points in the interior. It is also naturally isomorphic to the mapping class group of the surface obtained from a torus by deleting the interior of a closed, embedded disk.

The standard generating set for  $B_3$  is  $\Gamma = \{\sigma_1, \sigma_2\}$ . Each generator  $\sigma_i$  is a half-twist. We consider the pseudo-Anosov braid

$$f = \sigma_2 \sigma_1^{-1} \sigma_2 \sigma_1 \sigma_1 \sigma_1$$

We will use our algorithm to show that the stretch factor of  $f$  is

$$\lambda = \frac{5 + \sqrt{21}}{2}$$

and to obtain a description of the unstable foliation for  $f$  in terms of train tracks.

Our method in this section diverges from the main algorithm of this paper in that we do not make use of the explicit  $Q$  from Theorem 1.1. Instead, we use a guess-and-check approach. Specifically, we use the basic computation algorithm to compute  $f(c)$  and a triple  $(A, V_i, V_j)$  for the action of  $f$  at  $f(c)$ . Then we find an eigenvector  $v$  for  $A$  and check directly that it is a PL-eigenvector for  $f$  with eigenvalue  $\lambda > 1$ ; it follows that  $v$  represents the sink (if  $v$  were not a PL-eigenvector, we would proceed by computing the action of  $f$  at  $f^2(c)$ , etc.). This method always terminates, and in practice it can be more efficient than our main algorithm (since  $Q$  is so large). Our Theorem 1.1 guarantees that this method will terminate in quadratic time.

The discussion in this section has one other small difference from our general setup, namely, the fact that  $D_3$  has boundary; on the other hand, we could delete

the boundary and consider  $f$  as an element of  $\text{Mod}(S_{0,4})$  without changing the stretch factor or foliations. Note that

$$Q(S_{0,4}) = 2,464 + 96 + 1 = 2,561$$

which is much larger than the power of  $f$  required for the guess-and-check method, which in this case is 1.

**2.1. The ledger.** We start by setting up the ledger for the  $B_3$ -action on  $\text{MF}(D_3) \cong \mathbb{R}^2$ . Since we already defined the generating set  $\Gamma$ , it remains to choose a cell decomposition  $Y$  for  $\text{MF}(D_3)$  and then find the collection of subdivisions  $Y_\Gamma$  and the resulting collection of matrices  $\mathcal{A}$ .

*The cell decomposition.* The four train tracks  $\tau_1, \tau_2, \tau_3,$  and  $\tau_4$  shown in the left-hand side of Figure 2 give a cell decomposition  $Y$  of  $\text{MF}(D_3)$  as an integral cone complex. The top-dimensional cells are the  $V_i = V(\tau_i)$ , and the faces are given by the rays through four curves shown. For further explanation of this picture, see the book by Farb and the first author [11, Chapter 15] (the train tracks there are slightly different). Here we have another difference from our general setup, as the train tracks here are not Dehn–Thurston train tracks; on the other hand these train tracks are especially suited to the braid group.

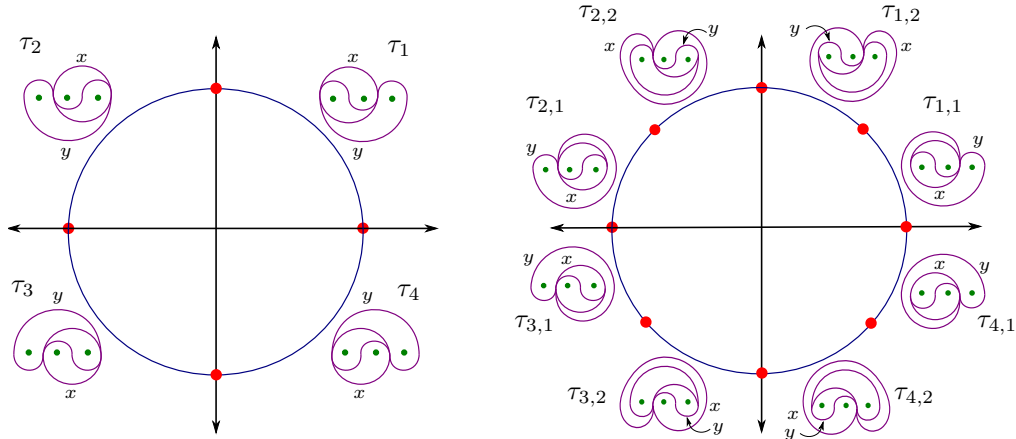


FIGURE 2. The integral cone complex structure  $Y$  on  $\text{MF}(D_3)$  and the common refinement of  $Y_\Gamma$

*The subdivisions and the matrices.* The eight train tracks  $\tau_{1,1}, \tau_{1,2}, \dots, \tau_{4,1}, \tau_{4,2}$  shown in the right-hand side of Figure 2 give the common refinement of the collection of subdivisions  $Y_\Gamma$  of  $Y$  associated to  $\Gamma$ . Indeed, for each  $f \in \Gamma^\pm$  and each  $\tau_{i,j}$  there is a  $k$  so that  $f(V_{i,j}) \subseteq V_k$  and  $f|_{V_{i,j}}$  is linear. We illustrate this claim in one case in Figure 3 (for an example of a similar computation, see again the book by Farb and the first author). The figure shows  $\tau_{1,2}, \sigma_1(\tau_{1,2})$ , and  $\tau_2$ , along with the weights corresponding to the  $xy$ -coordinates on  $V_{1,2}$ . We see from the figure that

$$\sigma_1(V_{1,2}) \subseteq V_2.$$

For each generator (or inverse of a generator)  $h$ , the subdivision  $Y_h$  lies between  $Y$  and the common subdivision. For instance, the subdivision  $Y_{\sigma_1}$  has top-dimensional cells  $V_{1,1}, V_{1,2}, V_2, V_3,$  and  $V_4$ . In each of the other three cases, the subdivision  $Y_h$  is again obtained by subdividing one cell of  $Y$ .



The calculations that allow us to check the inclusions  $f(V_{i,j}) \subseteq V_k$  also allow us to compute the collection of matrices  $\mathcal{A}$ . From the weights in Figure 3 we see that the matrix  $A_{1,2}^{\sigma_1}$  describing the action of  $\sigma_1$  on  $V_{1,2}$  is

$$A_{1,2}^{\sigma_1} = \begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix}$$

Since the coordinates on  $V_{1,2}$  are compatible with the coordinates on  $V_2$  (that is, the inclusion is given by the identity matrix), the matrix  $A_{1,2}^{\sigma_1}$  does (as required) describe the map  $\sigma_1|_{V_{1,2}}$  in terms of the  $V_1$ - and  $V_2$ -coordinates. As such, we can describe the action of  $\sigma_1$  on its linear piece  $V_{1,2}$  by the triple  $(A_{1,2}^{\sigma_1}, V_1, V_2)$ . The computations for the other cases are similar.

From the refinement given in Figure 2, we also obtain notation for each subdivision  $Y_{\sigma_i^{\pm 1}}$ . For example,  $Y_{\sigma_1}$  is the subdivision given by  $V_{1,1}, V_{1,2}, V_2, V_3,$  and  $V_4$ . While we will not list the subdivisions for the other 7 generators and inverses of generators, the subdivisions are implicit from the calculations below: if we write  $A_{1,2}^{\sigma_1}$  this signifies that  $V_{1,2}$  is a linear piece for  $\sigma_1$ , and if we write  $A_2^{\sigma_1}$  this signifies that  $V_2$  is a linear piece for  $\sigma_1$ .

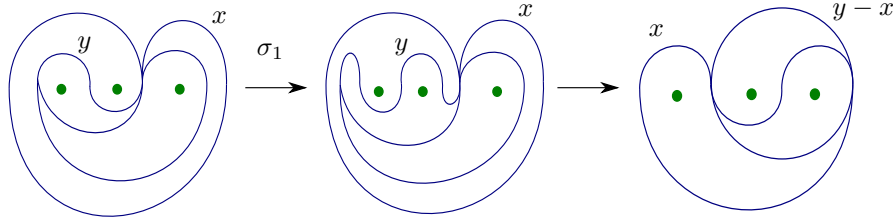


FIGURE 3. A part of the computation of the linear transformation  $V_{12} \rightarrow V_2$  induced by  $\sigma_1$

**2.2. Applying the algorithm.** As above, let  $f = \sigma_2\sigma_1^{-1}\sigma_2\sigma_1\sigma_1\sigma_1$ . Let  $c$  be the curve described by the vector  $(1, 2)$  in  $V_1$ -coordinates (again, contrary to our general setup, this curve does not lie on a vertex ray). We first use the basic computation algorithm to give a detailed description of how to compute  $f(c)$ , and then use the modified version of the main algorithm to find the stretch factor and foliations.

*Applying the basic computation algorithm.* Thinking of  $f$  as a word  $w$  in  $\Gamma$  we have  $f_1 = \sigma_1, f_2 = \sigma_1,$  etc. As in the description of the basic computation, we initialize by setting  $c_0 = c$ .

The following table gives the details of the steps of the basic computation for finding  $f(c)$ . In each row we give:

- (1) the step number  $\ell,$
- (2) the curve  $c_{\ell-1}$  from the previous step,
- (3) the cell  $V_{i,k}$  of the subdivision  $Y_\Gamma$  containing  $c_{\ell-1},$
- (4) the next letter  $f_\ell$  in the word  $w,$
- (5) the codomain  $V_j$  for  $f_\ell|_{V_{i,k}},$  and
- (6) the matrix  $A_{i,k}^\ell$  so that  $f_\ell|_{V_{i,k}}$  is described by  $(A_{i,k}^\ell, V_i, V_j).$

Here is the table:

$\ell$	$c_{\ell-1}$	$V_{i,k}$	$f_\ell$	$V_j$	$A_{i,k}^\ell$	$c_\ell$
1	$\begin{pmatrix} 1 \\ 2 \end{pmatrix}$	$V_{1,2}$	$\sigma_1$	$V_2$	$\begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 1 \end{pmatrix}$
2	$\begin{pmatrix} 1 \\ 1 \end{pmatrix}$	$V_2$	$\sigma_1$	$V_3$	$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 1 \end{pmatrix}$
3	$\begin{pmatrix} 1 \\ 1 \end{pmatrix}$	$V_3$	$\sigma_1$	$V_4$	$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 2 \end{pmatrix}$
4	$\begin{pmatrix} 1 \\ 2 \end{pmatrix}$	$V_4$	$\sigma_2$	$V_1$	$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$	$\begin{pmatrix} 2 \\ 1 \end{pmatrix}$
5	$\begin{pmatrix} 2 \\ 1 \end{pmatrix}$	$V_1$	$\sigma_1^{-1}$	$V_1$	$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 3 \\ 1 \end{pmatrix}$
6	$\begin{pmatrix} 3 \\ 1 \end{pmatrix}$	$V_1$	$\sigma_2$	$V_2$	$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 4 \end{pmatrix}$

(For  $\ell > 1$  we use  $V_i$  instead of  $V_{i,k}$  since the entire  $V_i$  is a linear piece for the corresponding  $f_\ell$ .) From the table, we see that

$$f(c) = \begin{pmatrix} 1 \\ 4 \end{pmatrix} \in V_2.$$

*An aside.* We can further see from the above computation that  $c$  lies in a linear piece for  $f$  described by a triple  $(A, V_1, V_2)$ , where  $A$  is given by the product of the above matrices:

$$\begin{aligned} A &= A_1^{\sigma_2} A_1^{\sigma_1^{-1}} A_4^{\sigma_2} A_3^{\sigma_1} A_2^{\sigma_1} A_{1,2}^{\sigma_1} \\ &= \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} -1 & 1 \\ -2 & 3 \end{pmatrix} \end{aligned}$$

We can check this by observing that

$$\begin{pmatrix} -1 & 1 \\ -2 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ 4 \end{pmatrix}$$

We will not use this description of action of  $f$  on the linear piece containing  $c$ ; we include it just to illustrate how the computations work.

*Iterative-guess-and-check algorithm.* We now describe in more detail the guess-and-check algorithm described at the start of the section. This method is the underpinning of some of the computer programs that predate our work. The idea of the algorithm is that, starting with  $k = 1$ , we “guess” that  $f^k(c)$  lies in a Nielsen–Thurston eigenregion for  $f$ , and then

- (1) compute a triple  $(A, V_i, V_j)$  for the action of  $f$  on a linear piece containing the curve  $f^k(c)$ ,
- (2) find pairs  $(\lambda, v)$  where  $\lambda > 1$ , where  $v$  lies in  $V_i \cap V_j$ , and where  $f(v) = \lambda v$  (if  $V_i = V_j$  such  $v$  can be computed as classical eigenvectors), and
- (3) check if one pair  $(\lambda, v)$  with  $\lambda > 1$  and  $v \geq 0$  is a PL-eigenvalue/eigenvector pair for  $f$ .

Since  $f$  has exactly one PL-eigenvalue/eigenvector pair with  $\lambda > 1$  (up to scale), the last step implies that  $\lambda$  is the stretch factor and  $v$  represents the unstable foliation for  $f$ . If either the second or third step fails (meaning no pairs are found in the second step, or it is found in the third step that they are not PL-eigenvectors), we return to the first step, replacing  $f^k(c)$  with  $f^{k+1}(c)$ . If  $f$  is pseudo-Anosov,

then this process eventually terminates, since  $f^k(c)$  must lie in a Nielsen–Thurston eigenregion for  $k$  large.

Again, the complexity of the iterative-guess-and-check algorithm is impossible to determine without an upper bound on the power of  $f$  required. Our Theorem 1.1 exactly gives such an upper bound, namely  $Q$ .

*Applying the iterative-guess-and-check algorithm.* We now apply the above algorithm to our specific  $f \in B_3$ . As we will see the algorithm terminates for  $k = 1$ . So this process terminates faster than if we were to apply the main algorithm, which would require computing  $f^Q(c)$  with  $Q = 2561$ .

(1) For the first step we apply the basic computation algorithm to the action of  $f$  on  $f(c)$  and find the resulting acting matrix:

$$\begin{aligned} A &= A_2^{\sigma_2} A_3^{\sigma_3^{-1}} A_4^{\sigma_4} A_4^{\sigma_4^{-1}} A_3^{\sigma_3} A_2^{\sigma_2} \\ &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix} \end{aligned}$$

This computation is done in the same way as the computation of  $f(c)$ . We do not include the corresponding table, but this can be inferred from the given product. During the computation, the curve  $f(c) \in V_2$  gets mapped to  $V_3$ , to  $V_4$ , to  $V_4$ , to  $V_1$ , again to  $V_1$ , and back  $V_2$ , in that order. In particular  $f(f(c))$  lies in  $V_2$ . This means that the result of the first step is the triple:

$$\left( \begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix}, V_2, V_2 \right)$$

(2) We now proceed to the second step. Since we have  $V_i = V_j$ , and we are guessing that  $f(f(c))$  lies in a Nielsen–Thurston eigenregion for  $f$ , we compute the eigenvalue  $\lambda > 1$  and corresponding eigenvector for the matrix  $A$  found in the first step. We find

$$\lambda = \frac{5 + \sqrt{21}}{2} \quad \text{and} \quad v = \begin{pmatrix} 1 + \sqrt{21} \\ 2 \end{pmatrix}.$$

(3) For the third and final step, we again use the basic computation algorithm to compute the action of  $f$  on a linear piece containing the eigenvector found in the last step. This time the computation turns out to be exactly the same as the computation in the first step, as the curves computed during the computation lie in exactly the same linear pieces for the six letters in the word for  $f$ . It follows that the pair  $(\lambda, v)$  found above is a PL-eigenvalue/eigenvector pair for  $f$ . As above, this means that  $\lambda$  is the stretch factor for  $f$  and the measured train track  $(\tau_2, v)$  represents the unstable foliation.

In order to find coordinates for the stable foliation for  $f$ , we can repeat the whole process, except with  $f$  replaced by  $f^{-1}$ . We find that the stable foliation is represented by  $(-3 + \sqrt{21}, 6)$  in  $V_1$ -coordinates.

### 3. SLOPE

The main goal of this section is to prove Proposition 3.1, which bounds the diameter of the horizontality of a curve or leaf of a foliation. We also introduce other tools related to slope and horizontality. At the end of the section we give our application of slope to translation lengths in the curve graph, Corollary 3.7.

In the statement of Proposition 3.1, the constant  $K = K(S_{g,p})$  is the number  $4(3g - 3 + p)$  from the introduction. Recall also that  $\mathcal{H}_f$  is the horizontality function coming from any singular Euclidean surface associated to a pseudo-Anosov  $f$ .

**Proposition 3.1.** *Let  $S = S_{g,p}$  and let  $f: S \rightarrow S$  be a pseudo-Anosov homeomorphism with stretch factor  $\lambda$  and a singular Euclidean structure  $X$ . If  $Z$  is a set of pairwise non-crossing saddle connections in  $\bar{X}$ , then*

$$\text{diam } \mathcal{H}_f(Z) \leq Q_{\text{curve}}(S).$$

In our applications of Proposition 3.1 we will take  $Z$  to be the set of saddle connections appearing in the (flat) geodesic representative  $c_X$  of a curve  $c$  or a leaf of a foliation in  $S$ . This makes sense because the set of saddle connections appearing in the geodesic representative  $c_X$  of a curve or a leaf are pairwise non-crossing. More generally, Minsky and the third-named author of this paper proved that if  $c$  and  $d$  are two disjoint curves/leaves, then the geodesic representatives  $c_X$  and  $d_X$  consist of pairwise non-crossing saddle connections; see [21, Lemma 4.3].

*Outline of the section.* We begin in Section 3.1 with some preliminary definitions: completed universal covers, rectangles, and rectangle flexibility. The latter, denoted  $\text{flex}(X)$ , describes the amount that rectangles in a singular Euclidean surface can be extended in the horizontal and vertical directions.

Next, in Section 3.2 we prove several lemmas that compare slopes of saddle connections in a fixed singular Euclidean surface. We first prove Lemma 3.2, a local slope disparity lemma, which bounds in terms of  $\text{flex}(X)$  the disparity between the slopes of saddle connections in a single rectangle (we obtain Lemma 3.2 as a consequence of Lemma 3.3, which compares the slope of a saddle connection to the slope of a rectangle containing it). We then promote Lemma 3.2 to Lemma 3.4, a global slope disparity lemma, which bounds in terms of  $\text{flex}(X)$  the disparity between the slopes of any non-crossing saddle connections in  $X$ .

In Section 3.3 we consider singular Euclidean surfaces together with associated pseudo-Anosov maps. In Lemma 3.6 we relate  $\text{flex}(X)$  to the stretch factor of a pseudo-Anosov map. We then turn to the proof of Proposition 3.1. Since the slope of a curve is the set of slopes of a collection of non-crossing saddle connections, Proposition 3.1 will follow from Lemmas 3.4 and 3.6.

Finally, in Section 3.4, apply the notion of horizontality to prove Corollary 3.7. This corollary bounds from below the stable translation length of the action of a pseudo-Anosov mapping class on the corresponding curve graph, extending the work of Gadre–Tsai.

**3.1. Preliminaries.** In this section, we introduce here the definitions of completed universal covers, rectangles, and rectangle flexibility, along with several related notions.

*Completed universal covers.* Let  $X$  be a singular Euclidean surface and let  $\bar{X}$  be the associated closed singular Euclidean surface, obtained from  $X$  by filling in the punctures. As in the work of Minsky and the third-named author, we denote by  $\hat{X}$  the metric completion of the universal cover of  $X$ . We refer to  $\hat{X}$  as the completed universal cover of  $X$ . There is an induced projection  $\hat{X} \rightarrow \bar{X}$ . This map can be regarded as a branched cover, with infinite-fold ramification over the points of  $\bar{X} \setminus X$ .

The universal cover  $\tilde{X}$  can be identified with  $\mathbb{H}^2$ , and as such the completed universal cover  $\hat{X}$  can be identified with a subset (but not a subspace!) of the visual compactification of  $\mathbb{H}^2$ . The space  $\tilde{X}$  is a CAT(0) space. In particular, given a pair of points in  $\tilde{X}$ , there is a unique geodesic connecting them. This geodesic may pass through arbitrarily many points of  $\hat{X} \setminus \tilde{X}$ .

*Rectangles and saddle connections.* Let  $X$  be a singular Euclidean surface with horizontal and vertical foliations  $\mathcal{F}_h$  and  $\mathcal{F}_v$ . A rectangle in  $X$  or  $\bar{X}$  is defined to be an isometric immersion  $R: E \rightarrow \bar{X}$ , where

- (1)  $E$  is a Euclidean rectangle,
- (2) each horizontal/vertical segment in  $E$  maps to a horizontal/vertical segment, and
- (3)  $R^{-1}(\Sigma) \subseteq \partial E$ .

While  $R$  is an immersion, it is always possible to lift  $R$  to  $\hat{X}$  and any such lift will be an embedding. In all of our figures, the rectangles will be shown as embedded; we can (safely) imagine that these pictures describe lifts to  $\hat{X}$ .

By a diagonal of  $R$  we mean the  $R$ -image of a diagonal of  $E$ . When we say that a point  $x \in X$  lies in the boundary or interior of  $R$ , we mean that  $R^{-1}(x)$  lies in the boundary or interior of  $E$ , respectively. The length and width of  $R$  are defined to be the length and width of  $E$ .

We say that a saddle connection  $\sigma$  in  $\bar{X}$  is rectangle spanning if it is the diagonal of a rectangle. In this case the rectangle is denoted  $R_\sigma$  and is called the spanning rectangle for  $\sigma$ . The length and width of a rectangle-spanning saddle connection are defined to be the length and width of its spanning rectangle.

Minsky and the third-named author proved that if  $X$  a singular Euclidean surface associated to a pseudo-Anosov mapping class  $f$ , the triangulations of  $X$  by rectangle-spanning saddle connections are exactly the sections of Agol's veering triangulation [21, Section 3].

*Slopes of rectangles and flexibility.* Let  $X$  be a singular Euclidean surface. As in the introduction, we define the slope of a rectangle  $R \subseteq X$  as  $\text{slope}(\delta)$ , where  $\delta$  is either diagonal of  $R$  (the diagonals have equal slope). Also, we say that a rectangle  $\hat{R}$  is an extension of  $R$  if it is obtained by extending  $R$  in one direction (up, down, left, or right). Finally, we define the rectangle flexibility of  $X$  as

$$\text{flex}(X) = \sup_{(\sigma, \hat{R}_\sigma)} \max \left\{ \frac{\text{slope}(\sigma)}{\text{slope}(\hat{R}_\sigma)}, \frac{\text{slope}(\hat{R}_\sigma)}{\text{slope}(\sigma)} \right\},$$

where  $\sigma$  is a rectangle-spanning saddle connection in  $X$  and  $\hat{R}_\sigma$  is an extension of the spanning rectangle  $R_\sigma$ .

*Finite flexibility versus the existence of horizontal and vertical saddle connections.* For a singular Euclidean surface  $X$  we have the following equivalence:

$$\text{flex}(X) < \infty \iff \bar{X} \text{ has no horizontal or vertical saddle connection}$$

Much of our work in this section relies on the work of Minsky and the third-named author of this paper [21]. In their work on singular Euclidean surfaces, they assume the surfaces do not have any horizontal or vertical saddle connections. As a proxy for this assumption, we will assume the condition that  $X$  has finite flexibility. It is a straightforward consequence of the definition of flexibility that this implies the non-existence of vertical and horizontal saddle connections. The other direction of the implication is also true, but we will not require it.

A singular Euclidean structure for a pseudo-Anosov homeomorphism has no vertical or horizontal saddle connections. Such a structure also has finite flexibility by Lemma 3.6 (or by the reverse implication above). We will use this without mention below in order to apply our lemmas about singular Euclidean structures with finite flexibility to singular Euclidean structures associated to pseudo-Anosov homeomorphisms.

**3.2. Slope disparity lemmas.** The main goals of this section are to prove Lemmas 3.2 and 3.4, the local and global slope disparity lemmas. In what follows, we will write  $F = F(X)$  for the quantity  $2 \text{flex}(X) - 1$ ; this quantity will appear often in what follows.

*Local slope disparity.* The next lemma bounds the disparity between slopes of saddle connections in a given rectangle.

**Lemma 3.2.** *Let  $X$  be a singular Euclidean surface, let  $F = F(X)$ , let  $R$  be a rectangle in  $X$ , and let  $\sigma_0$  and  $\sigma_1$  be saddle connections contained in  $R$ . Then*

$$\text{slope}(\sigma_0) \leq F^2 \text{slope}(\sigma_1).$$

Lemma 3.2 is an immediate consequence of the following lemma, which compares the slope of a rectangle to the slopes of saddle connections contained within.

**Lemma 3.3.** *Let  $X$  be a singular Euclidean surface, let  $F = F(X)$ , and let  $R$  be a rectangle in  $X$  and  $\sigma$  a saddle connection in  $R$ . Then*

$$\frac{1}{F} \leq \frac{\text{slope}(\sigma)}{\text{slope}(R)} \leq F.$$

*Proof.* We prove the second inequality. The first is proved with a symmetric argument. Let  $R_\sigma^h$  be the maximal rectangle obtained by extending the spanning rectangle  $R_\sigma$  horizontally in both directions. Note that  $\sigma$  connects the horizontal edges of  $R_\sigma^h$ .

As  $\text{slope}(R_\sigma^h) \leq \text{slope}(R)$  and  $F = 2 \text{flex}(X) - 1$ , it suffices to prove that

$$\frac{\text{slope}(\sigma)}{\text{slope}(R_\sigma^h)} \leq 2 \text{flex}(X) - 1.$$

Since ratios of slopes are invariant under horizontal and vertical scaling, we may assume for simplicity that

$$\text{height}(R_\sigma) = \text{width}(R_\sigma) = 1.$$

Let  $R_\sigma^-$  and  $R_\sigma^+$  be the maximal extensions of  $R_\sigma$  to the left and right. By definition, we have

$$\text{width}(R_\sigma^\pm) \leq \text{flex}(X).$$

If  $E$ ,  $E^h$ ,  $E^-$ , and  $E^+$  are the domains of  $R_\sigma$ ,  $R_\sigma^h$ ,  $R_\sigma^-$ , and  $R_\sigma^+$ , respectively, then  $E^- \cup E^+ = E^h$  and  $E^- \cap E^+ = E$ . Therefore

$$\begin{aligned} \frac{\text{slope}(\sigma)}{\text{slope}(R_\sigma^h)} &= \frac{1}{\text{slope}(R_\sigma^h)} = \text{width}(R_\sigma^h) \\ &= \text{width}(R_\sigma^+) + \text{width}(R_\sigma^-) - \text{width}(R_\sigma) \leq 2 \text{flex}(X) - 1, \end{aligned}$$

as desired.  $\square$

*Global slope disparity.* Our next goal is to prove the following lemma, which bounds the disparity between slopes of non-crossing saddle connections.

**Lemma 3.4.** *Let  $X$  be a singular Euclidean surface and let  $F = F(X)$ . If  $\sigma_0$  and  $\sigma_1$  are non-crossing saddle connections in  $X$ , then*

$$\text{slope}(\sigma_0) \leq F^{2K} \text{slope}(\sigma_1).$$

Our proof of Lemma 3.4 requires a subordinate lemma, Lemma 3.5 below. Before stating that lemma, we recall some preliminaries from the work of Minsky and the third-named author, in particular the notion of a rectangle hull; see [21, Section 4.1]. We also introduce the notion of a rectangle associated to a point.

Let  $X$  be a singular Euclidean surface with finite flexibility. For any saddle connection  $\sigma$  of  $X$ , let  $\mathcal{R}(\sigma)$  be the collection of rectangles in  $X$  that are maximal with respect to the property that their diagonals lie along  $\sigma$ . We refer to  $\mathcal{R}(\sigma)$  as the rectangle casing for  $\sigma$ .

The rectangle hull  $\mathbf{r}(\sigma)$  is defined to be the collection of saddle connections contained in some rectangle of  $\mathcal{R}(\sigma)$ . For example, the rectangle hull of the saddle connection in Figure 4 contains the four saddle connections that join the singularities in the boundary of  $R^+$  or the boundary of  $R^-$ . Every rectangle hull is connected (the rectangles in  $\mathcal{R}(\sigma)$  are ordered along  $\sigma$  and consecutive rectangles share a singularity). If  $\sigma_0$  and  $\sigma_1$  are non-crossing saddle connections, then the saddle connections of  $\mathbf{r}(\sigma_0) \cup \mathbf{r}(\sigma_1)$  are also non-crossing [21, Lemma 4.1].

If  $C$  is a singular point in a rectangle  $R \in \mathcal{R}(\sigma)$  and does not lie on  $\sigma$ , then we define the associated rectangle  $R_C$  as the rectangle in  $R$  with corner  $C$  and diagonal along  $\sigma$ . In Figure 4 we show a representative configuration.

Let  $R_C^-$  and  $R_C^+$  be the rectangles of  $\mathcal{R}(\sigma)$  obtained by sliding one of the points of  $R_C \cap \sigma$  along  $\sigma$ . The rectangle  $R_C$  is contained in both of these. Also it is possible that one or both is equal to  $R_C$ .

The set  $\mathcal{R}(\sigma)$  is finite, because the diagonals of the elements of  $\mathcal{R}(\sigma)$  form a pairwise non-nested cover of  $\sigma$  by closed intervals. Also, each element of  $\mathcal{R}(\sigma)$  contains at most one singularity on each edge, since we have assumed that  $X$  has no horizontal or vertical saddle connections. In particular, there are finitely many  $R_C$  as above.

**Lemma 3.5.** *Let  $X$  be a singular Euclidean surface with finite flexibility. For any saddle connection  $\sigma$ , there are rectangle-spanning saddle connections  $\sigma^+$ ,  $\sigma^-$  in the rectangle hull  $\mathbf{r}(\sigma)$  such that*

$$\text{slope}(\sigma^-) \leq \text{slope}(\sigma) \leq \text{slope}(\sigma^+).$$



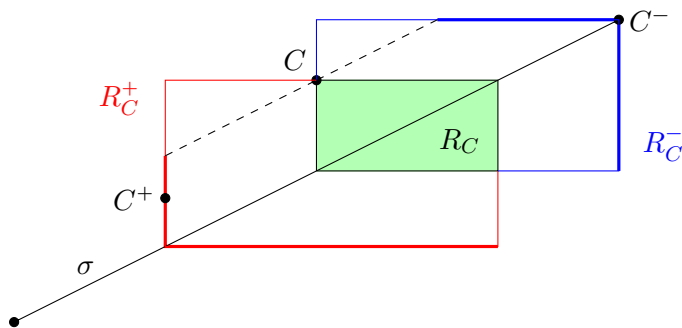


FIGURE 4. The rectangles  $R_C$ ,  $R_C^+$ , and  $R_C^-$  from the proof of Lemma 3.5

*Proof.* As above, there are finitely many rectangles  $R_C$  associated to singularities in the rectangles of  $\mathcal{R}(\sigma)$ . Thus there is such an  $R_C$  of largest area.

Without loss of generality, we assume that  $\sigma$  passes through the bottom-left and top-right corners of  $R_C$  and that  $C$  lies above (or at an endpoint of)  $\sigma$  (if needed, we precompose  $R$  with a symmetry of its domain rectangle  $E$ ). Let  $R_C^+$  and  $R_C^-$  denote the elements of  $\mathcal{R}(\sigma)$  that contain  $R_C$  and are obtained by sliding, respectively, the bottom-left and top-right corners of  $R_C$  along  $\sigma$ . As above, it is possible for one or both of these two to equal  $R_C$ . See Figure 4 for an illustration.

There must be a singularity  $C^+$  on the left or bottom boundary of  $R_C^+$  and also a singularity  $C^-$  on the top or right boundary of  $R_C^-$ , for otherwise,  $R_C^+$  and  $R_C^-$  would not be maximal (as demanded by the definition of the rectangle hull).

Let  $\ell$  be the line segment in  $R_C^+ \cup R_C^-$  that passes through  $C$  and is parallel to  $\sigma$ . The singularities  $C^+$  and  $C^-$  lie on the same side of  $\ell$  as  $\sigma$ . Indeed, if not, there would be a translate of  $R_C^\pm$  (in the direction of  $\sigma$ ) that strictly contains  $R_C$ , violating the maximality of the latter.

Let  $\sigma^+$  and  $\sigma^-$  be the saddle connections connecting  $C^+$  and  $C^-$  to  $C$  within  $R_C^+$  and  $R_C^-$ . We have by definition that  $\sigma^+$  and  $\sigma^-$  lie in  $\mathbf{r}(\sigma)$ . By the previous paragraph they satisfy  $\text{slope}(\sigma^-) \leq \text{slope}(\sigma) \leq \text{slope}(\sigma^+)$ , as desired.  $\square$

We now turn towards the proof of Lemma 3.4. In the proof, we use the fact that  $K(S)$ , our upper bound for the number of separatrices in a foliation on  $S$ , is also an upper bound for the number of triangles in a triangulation of  $S$  by saddle connections. Indeed, the map associating each separatrix to the triangle it initially enters is bijective.

*Proof of Lemma 3.4.* Since the lemma is vacuously true for surfaces of infinite flexibility, we may assume that  $X$  has finite flexibility. Hence we may apply the theory of rectangle hulls described above.

We proceed in two steps, first proving the lemma in the case of rectangle-spanning saddle connections and then in the general case.

To this end, first suppose that  $\sigma_0$  and  $\sigma_1$  are non-crossing rectangle-spanning saddle connections in  $X$ . Minsky and the third-named author proved that any collection of non-crossing, rectangle-spanning saddle connections can be extended to a triangulation by rectangle-spanning saddle connections [21, Lemma 3.2]. Consider then such triangulation of  $X$ . As above, the number of triangles is bounded above

by  $K = K(S)$ . Thus, any sequence of distinct triangles  $(T_i)$  has length at most  $K$ . It further follows that there is a sequence of edges of the triangulation  $e_0 = \sigma_0, e_2, \dots, e_n = \sigma_1$  with  $n \leq K$  and with each consecutive pair contained in a triangle. Since each triangle is contained in a rectangle, the special case now follows from Lemma 3.2.

We now proceed to the general case. Let  $\sigma_0$  and  $\sigma_1$  be arbitrary non-crossing saddle connections. For  $i = 0, 1$  let  $\sigma_i^\pm$  be the rectangle-spanning saddle connections given by Lemma 3.5. Minsky and Taylor proved that if two saddle connections are non-crossing then the elements of their rectangle hulls are pairwise non-crossing [21, Lemma 4.1]. Therefore, by Lemma 3.5 and the special case of this lemma, we have

$$\frac{\text{slope}(\sigma_0)}{\text{slope}(\sigma_1)} \leq \frac{\text{slope}(\sigma_0^+)}{\text{slope}(\sigma_1^-)} \leq F^{2K},$$

as desired. □

**3.3. Slopes and stretch factors.** In this section we incorporate the slope disparity lemmas together with basic properties of pseudo-Anosov maps in order to prove Proposition 3.1. We begin with a comparison of rectangle flexibilities and stretch factors.

*Flexibility versus stretch factor.* The following lemma bounds the rectangle flexibility of a singular Euclidean surface corresponding to a pseudo-Anosov map in terms of the stretch factor.

**Lemma 3.6.** *Let  $S = S_{g,p}$  and let  $f: S \rightarrow S$  be pseudo-Anosov with stretch factor  $\lambda$  and singular Euclidean structure  $X$ . Then*

$$\text{flex}(X) \leq \lambda^K.$$

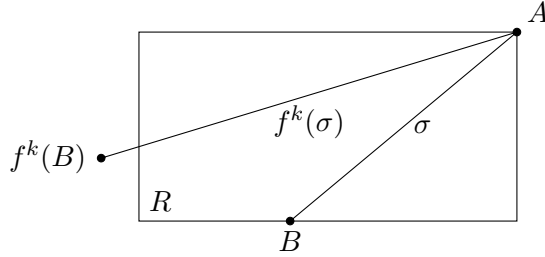


FIGURE 5. The saddle connections  $\sigma$  and  $f^k(\sigma)$  from the proof of Lemma 3.6

*Proof.* Let  $\hat{R}_\sigma$  be an extension of a spanning rectangle for a saddle connection  $\sigma$  in  $X$ . And let us denote  $\hat{R}_\sigma$  by  $R$ . By symmetry, we may assume that  $R$  is a horizontal extension, as in Figure 5 (for the other case we replace  $f$  with  $f^{-1}$ ). Say that the endpoints of  $\sigma$  are the singularities  $A$  and  $B$ . It must be the case that one of these, say  $A$ , is at a corner of  $R$ . The singularity  $B$  must then lie on a horizontal edge of  $R$ . We need to show two inequalities:

$$\frac{\text{slope}(R)}{\text{slope}(\sigma)} \leq \lambda^K \quad \text{and} \quad \frac{\text{slope}(\sigma)}{\text{slope}(R)} \leq \lambda^K.$$

The first inequality is immediate since  $\text{slope}(R) \leq \text{slope}(\sigma)$  and  $\lambda > 1$ . We proceed to the second inequality. Choose  $k \leq K$  so that  $f^k$  fixes  $A$  as well as the separatrices of the horizontal and vertical foliations at  $A$ . Since  $R$  contains no singularities,  $f^k(B)$  cannot lie in the interior of  $R$ . Thus

$$\text{width}(R) \leq \text{width}(f^k(\sigma)).$$

On the other hand, we have

$$\text{height}(R) = \text{height}(\sigma) = \lambda^k \text{height}(f^k(\sigma)).$$

Dividing the first and last expressions here by the expressions in the previous inequality, we obtain

$$\text{slope}(R) \geq \lambda^k \text{slope}(f^k(\sigma)).$$

Since  $\text{slope}(\sigma) = \lambda^{2k} \text{slope}(f^k(\sigma))$ , we have

$$\frac{\text{slope}(\sigma)}{\text{slope}(R)} \leq \frac{\lambda^{2k} \text{slope}(f^k(\sigma))}{\lambda^k \text{slope}(f^k(\sigma))} = \lambda^k \leq \lambda^K,$$

as desired.  $\square$

*Penner's lower bound.* Penner [25] proved for any pseudo-Anosov  $f \in \text{Mod}(S)$  that

$$2 \leq \lambda^K$$

(he stated the equivalent formulation  $\log \lambda(f) \geq (\log 2)/K$ ). In the proof of Proposition 3.1 we will use the inequality  $4^K \leq \lambda^{2K^2}$ , obtained from Penner's inequality by squaring both sides and raising to the  $K$ th power.

*Completing the proof.* We are now ready for the proof of Proposition 3.1, which gives an upper bound on the diameter of a set of non-crossing saddle connections.

*Proof of Proposition 3.1.* As in the statement of the proposition,  $S$  is a surface,  $f: S \rightarrow S$  is pseudo-Anosov with stretch factor  $\lambda$ ,  $X$  is an associated singular Euclidean surface, and  $Z$  is a collection of non-crossing saddle connections. Also recall that  $Q_{\text{curve}} = Q_{\text{curve}}(S) = 2K^2$ . Let  $F = F(X)$  be the flexibility. The conclusion of the proposition is equivalent to the inequality

$$\frac{\max \text{slope}(Z)}{\min \text{slope}(Z)} \leq (\lambda^2)^{Q_{\text{curve}}}.$$

We proceed to the proof of this inequality. Let  $\sigma_0$  and  $\sigma_1$  be any two saddle connections of  $Z$ . Combining Lemma 3.4, Lemma 3.6, the fact that  $2\lambda^K - 1 < 2\lambda^K$ , and Penner's theorem mentioned above, we have

$$\begin{aligned} \text{slope}(\sigma_0) &\leq \text{slope}(\sigma_1) F^{2K} = \text{slope}(\sigma_1) (2 \text{flex}(X) - 1)^{2K} \\ &\leq \text{slope}(\sigma_1) (2\lambda^K - 1)^{2K} < \text{slope}(\sigma_1) 4^K \lambda^{2K^2} \leq \text{slope}(\sigma_1) (\lambda^2)^{2K^2}, \end{aligned}$$

as required.  $\square$

**3.4. Application to stable translation lengths in the curve graph.** As above, Proposition 3.1 gives an upper bound of  $Q_{\text{curve}}$  to the diameter of the horizontality of a collection of pairwise non-crossing saddle connections. This proposition has the following corollary. In the statement,  $\mathcal{C}(S)$  is the curve graph for  $S$  and  $\tau_{\mathcal{C}(S)}(f)$  is the stable translation length for the action of a pseudo-Anosov  $f$  on  $\mathcal{C}(S)$ . The second statement of the corollary is a (slight) quantitative strengthening of a special case of a result of Gadre–Tsai [15, Theorem 1.1].

**Corollary 3.7.** *Let  $S = S_{g,p}$  and let  $f : S \rightarrow S$  be pseudo-Anosov. The horizontality function defines a  $Q_{\text{curve}}$ -coarse Lipschitz map*

$$\mathcal{H}_f : \mathcal{C}(S) \rightarrow \mathbb{R}.$$

Hence, for any pseudo-Anosov  $f \in \text{Mod}(S)$ , we have

$$\tau_{\mathcal{C}(S)}(f) \geq \frac{1}{Q_{\text{curve}}}.$$

For closed surfaces, Gadre–Tsai proved that for  $g \geq 2$  and any pseudo-Anosov  $f \in \text{Mod}(S_g)$  we have

$$\tau_{\mathcal{C}(S_g)}(f) \geq \frac{1}{162(2g-2)^2 + 30(2g-2)}.$$

Since  $Q_{\text{curve}} = 2K^2 < 162(2g-2)^2 + 30(2g-2)$  for  $g > 1$ , our Corollary 3.7 is a strengthening of this. For surfaces with punctures, Gadre–Tsai give a lower bound that is stronger than the one given by our corollary.

#### 4. APPROXIMATING RECTANGLES

The goal of this section is to prove Lemma 4.1. This is the most technical result in the paper. Let  $X$  be a singular Euclidean surface and let  $\hat{X}$  be its completed universal cover, as in Section 3. What Lemma 4.1 tells us is that if  $\gamma_1$  and  $\gamma_2$  are geodesics in  $\tilde{X}$  that cross each other and also cross a singularity-free horizontal leaf  $\ell$ , then there is a maximal rectangle  $R$  in  $\hat{X}$  with the following properties:

- $\gamma_1$  and  $\gamma_2$  cross the horizontal sides of  $R$ , and
- any saddle connection in  $X$  with sufficiently small slope has a path lift that crosses the vertical sides of  $R$ .

Before stating Lemma 4.1, we introduce the notion of a  $k$ -approximating rectangle.

*Approximating rectangles.* As above let  $X$  be a singular Euclidean surface with  $\text{flex}(X) < \infty$  and let  $\hat{X}$  be its completed universal cover. Let  $\gamma_1$  and  $\gamma_2$  be intersecting geodesics in  $\hat{X}$  whose endpoints (if any) are singular points of  $\hat{X}$  (all points of  $\hat{X} \setminus \tilde{X}$  are singular).

Let  $\ell$  be a bi-infinite horizontal leaf in  $\hat{X}$  that intersects both  $\gamma_1$  and  $\gamma_2$ . We say a rectangle  $R$  in  $\hat{X}$  is a  $k$ -approximating rectangle for the triple  $(\gamma_1, \gamma_2, \ell)$  if

- $R$  is maximal (i.e. it contains a singularity in each edge),
- $\ell$  contains an interior horizontal leaf of  $R$ ,
- $\gamma_1$  and  $\gamma_2$  connect the horizontal edges of  $R$ , and
- $k \text{slope}(R) \geq \min \{ \text{slope}(\gamma_1) \cup \text{slope}(\gamma_2) \}$ .

As in Section 3, we write  $F(X)$  for the quantity  $2\text{flex}(X) - 1$  when  $X$  is a singular Euclidean surface. Also, in what follows, by a lift of a saddle connection will mean a path lift.

**Lemma 4.1.** *Let  $X$  be a singular Euclidean surface with  $\text{flex}(X) < \infty$ . Let  $\gamma_1$  and  $\gamma_2$  be intersecting geodesics in  $\widehat{X}$  whose endpoints (if any) are singular points of  $\widehat{X}$ . Let  $\ell$  be a singularity-free horizontal leaf in  $\widehat{X}$  that intersects both  $\gamma_1$  and  $\gamma_2$ . Let  $F = 2\text{flex}(X) - 1$  and let  $K = K(X)$  be the separatrix complexity. Then the following hold.*

- (1) *There exists an  $F^5$ -approximating rectangle  $R$  for  $(\gamma_1, \gamma_2, \ell)$ .*
- (2) *If  $R$  is an  $F^5$ -approximating rectangle for  $(\gamma_1, \gamma_2, \ell)$  and  $Z$  is a collection of pairwise non-crossing saddle connections in  $X$  with*

$$\min \text{slope}(Z) < F^{-(2K+6)} \text{slope}(\gamma_i) \text{ for } i \in \{1, 2\}$$

*then there is a  $\sigma \in Z$  and a lift  $\tilde{\sigma}$  that crosses the vertical edges of  $R$ .*

**4.1. The existence of approximating rectangles.** For the proof of the first statement of Lemma 4.1—on the existence of approximating rectangles—we introduce three ingredients: a basic fact about triangles in singular Euclidean surfaces, developing maps, and a graph of rectangles and rectangle-spanning saddle connections.

*Triangles in singular Euclidean surfaces.* In the proof of Lemma 4.1, we will use that fact that a triangle in a (simply connected) singular Euclidean surface—a region bounded by three geodesics—has no singularities in its interior. This follows readily from the Gauss–Bonnet formula for singular Euclidean surfaces with geodesic boundary, namely:

$$2\pi\chi(Y) = -\pi \left( \sum_{p_k} \text{deg}(p_k) \right) + \sum_i (\pi - \theta_i),$$

where the  $\theta_i$  are the interior angles on the boundary and each  $\text{deg}(p_k) \geq 1$  is the degree of the interior singularity  $p_k$  (the degree is two less than the number of prongs); see [27, Theorem 3.3].

*Developing maps.* Let  $X$  be a singular Euclidean surface. By a disk in  $\widehat{X}$  we mean an isometric embedding of a closed polygonal disk in the Euclidean plane. Examples of disks are (the images of) rectangles. It follows from the Gauss–Bonnet formula that a triangle (as in the previous paragraph) is a disk in this sense.

Suppose we have a collection of disks  $D_0, D_1, \dots$  in  $\widehat{X}$ . Suppose that each  $D_i$  with  $i > 0$  has the property that its intersection with  $D_0 \cup \dots \cup D_{i-1}$  is connected and contains a nonempty open set. In this case there is a developing map from  $\cup D_i$  to the Euclidean plane that respects horizontal and vertical directions and measures. This map is defined inductively. First we make a choice of embedding of  $D_0$  in the plane; this choice is unique up to translation, rotation by  $\pi$ , and reflection through a horizontal or vertical line. Then each subsequent  $D_i$  is mapped to the plane in a way that agrees with the intersection with  $D_0 \cup \dots \cup D_{i-1}$ . Given a choice of embedding of  $D_0$ , the rest of the developing map is determined.

In the proof of Lemma 4.1, we will use the (almost) uniqueness of the developing map as a way of relating certain configurations of triangles and rectangles in a

singular-Euclidean surface to a configuration of disks in the Euclidean plane. Once we fix a developing map it makes sense to use words like “above” and “to the left” when describing features of the configuration of disks.

*The graph of rectangles and rectangle-spanning saddle connections.* Let  $X$  be a singular Euclidean surface. We define  $\mathcal{G}_X$  to be the graph whose vertices are the rectangles in  $X$  and the rectangle-spanning saddle connections in  $X$ . We connect a rectangle-spanning saddle connection  $\sigma$  to a rectangle  $R$  if  $\sigma$  is contained in  $R$ . It follows immediately from Lemma 3.3 that if  $v_0, \dots, v_n$  is a path in  $\mathcal{G}_X$  then

$$F^n \text{slope}_X(v_0) \geq \text{slope}_X(v_n),$$

In the proof of Lemma 4.1 we will bound the slope of the approximating rectangle by constructing a path in the graph  $\mathcal{G}_{\hat{X}}$  associated to  $\hat{X}$ . We will take  $v_0$  to be the approximating rectangle  $R$  and  $v_n$  to be a saddle connection of  $\gamma_1$  or  $\gamma_2$  (or a saddle connection with slope greater than some such saddle connection).

*Proof of Lemma 4.1(1).* We break the proof into two main parts. In the first part we construct a rectangle  $R$  satisfying the first three conditions of the definition of an approximating rectangle, and in the second part we show that  $R$  satisfies the slope condition. We begin with some setup. Throughout, the reader may refer to Figure 6.

*Setup.* Denote the unique points of intersection  $\ell \cap \gamma_1$  and  $\ell \cap \gamma_2$  by  $x_1$  and  $x_2$  (these points are unique because two geodesics must intersect in a connected union of saddle connections and because  $\ell$  contains no singularities). The geodesics  $\gamma_1$ ,  $\gamma_2$ , and  $\ell$  bound a triangle  $\Delta$ . As above the interior of  $\Delta$  is free of singularities. We allow for the case where  $x_1 = x_2$ ; in this case  $\Delta$  is degenerate, namely, it is the single point  $x_1$ . In what follows, we assume that we are in the generic case where  $x_1 \neq x_2$ , and so  $\Delta$  has nonempty interior.

As per Figure 6 we may choose a developing map for  $\Delta$  so that  $\Delta$  lies above  $\ell$ . We may also assume that  $\gamma_1$  and  $\gamma_2$  lie to the left and right of  $\Delta$ , respectively. These two choices determine the developing map up to translation.

Below, we will define rectangles  $R_1, \dots, R_5$ , and at each stage it will be true that the union satisfies the criterion for the existence of a developing map, which is at each stage uniquely defined as an extension of the developing map defined on  $\Delta$ . The images of the  $R_i$  and  $\Delta$  under the developing map (at least in a generic case) are shown in Figure 6.

*Part 1: Constructing  $R$ .* In order to construct  $R$ , we first construct three intermediate rectangles,  $R_1$ ,  $R_2$ , and  $R_3$ .

For each  $i \in \{1, 2\}$ , we define  $R_i$  to be a rectangle in  $\mathcal{R}(\gamma_i)$  satisfying the property that  $R_i$  contains singularities on either side of  $\ell$ . To explain why such rectangles exist, we fix some  $i \in \{1, 2\}$ . Since  $\ell$  intersects  $\gamma_i$ , it crosses some saddle connection  $\sigma_i$  of  $\gamma_i$ . As  $\ell$  contains no singular points, the two singularities of  $\sigma_i$  lie on opposite sides of  $\gamma_i$ . The rectangle hull  $\mathbf{r}(\sigma_i)$  is a connected graph in  $\hat{X}$  containing the two singular points of  $\sigma_i$ . Therefore, there is a saddle connection  $\rho_i$  of  $\mathbf{r}(\sigma_i)$  that connects singular points on opposite sides of  $\ell$ . The rectangle  $R_i \in \mathcal{R}(\sigma_i)$  associated to  $\rho_i$  satisfies the desired property.

We note that Figure 6 does not accurately reflect the case where  $R_1 \cap R_2 \neq \emptyset$ . But our arguments still apply in this case.

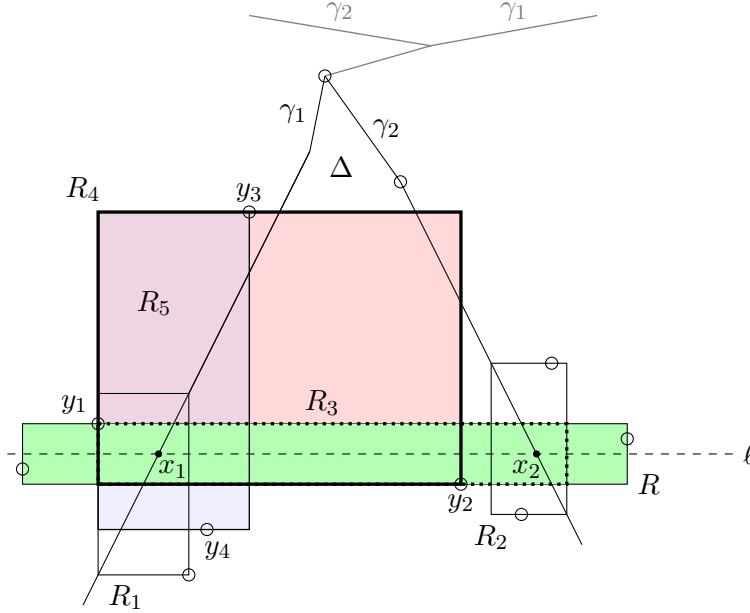


FIGURE 6. The rectangles and triangle in the proof of Lemma 4.1; the outline of  $R_3$  is dotted, while the outline of  $R_4$  is bold

We now proceed to construct the rectangle  $R_3$ . Let  $\ell_0$  be the minimal segment of  $\ell$  containing  $R_1 \cap \ell$  and  $R_2 \cap \ell$ . Then let  $R_3$  be the rectangle that is minimal with respect to the following properties:

- $\ell_0$  is a horizontal leaf of  $R_3$  and
- $R_3$  contains singularities on both its top and bottom edges.

The rectangle  $R_3$  has a singularity on both of its horizontal edges (possibly at the corners), but not in the interiors of its vertical edges. Because  $R_1$  and  $R_2$  have singularities on both sides of  $\ell$ , the endpoints of both vertical edges of  $R_3$  are contained in  $R_1 \cup R_2$ , as in Figure 6.

Finally, let  $R$  be the rectangle that is maximal with respect to the property that the horizontal edges of  $R_3$  are contained in  $R$ . In other words,  $R$  is the maximal horizontal extension of  $R_3$ .

To complete the first part of the proof, we check that  $R$  satisfies the first three conditions in the definition of an approximating rectangle.

- Since  $R_3$  contains singularities on both of its horizontal edges and  $R$  is a maximal horizontal extension of  $R_3$ , we have that  $R$  is maximal.
- Since  $R_3$  contains a segment of  $\ell$  in its interior and  $R$  contains  $R_3$ , we have that  $R$  contains a segment of  $\ell$  in its interior.
- Since  $R_1$  and  $R_2$  both contain singularities on both sides of  $\ell$ , the vertical edges of  $R_3$  are contained in the vertical edges of  $R_1$  and  $R_2$ . Then, since the diagonals of  $R_1$  and  $R_2$  are contained in  $\gamma_1$  and  $\gamma_2$ , respectively, it follows that  $\gamma_1$  and  $\gamma_2$  cross the horizontal edges of  $R_3$ , hence  $R$ , as desired.

This completes the first part of the proof.



*Part 2: Estimating the slope.* Our strategy for estimating the slope of  $R$ , and hence verifying the fourth condition of a  $k$ -approximating rectangle, is to find a path in  $\mathcal{G}_{\widehat{X}}$  from  $R$  to a saddle connection that has slope greater than or equal to the slope of some saddle connection of  $\gamma_1$  or  $\gamma_2$ . Since the goal is to show that  $R$  is an  $F^5$ -approximating rectangle, it suffices (as above) to find such a path whose length is bounded above by 5. To this end, we start by defining singular points  $y_1, y_2, y_3$ , and  $y_4$  and rectangles  $R_4$  and  $R_5$ ; each of the saddle connections we use in our path will be of the form  $y_i y_j$ .

- Let  $y_1$  and  $y_2$  be the singularities on the horizontal edges of  $R_3$  that lie above and below  $\ell$ , respectively.

These exist by the definition of  $R_3$ , and they are unique since  $\widehat{X}$  has no horizontal saddle connections.

Since  $\Delta$ ,  $R_1$ , and  $R_2$  have no singularities in their interiors, it follows that  $y_1$  lies in the boundary of  $R_1$  or  $R_2$ . Up to relabeling the  $\gamma_i$ , we may assume that it lies in the boundary of  $R_1$ , as in Figure 6. (Assuming  $R_1 \neq R_2$  we have no further flexibility in the choice of developing map.) Because  $\Delta$  has no singularities in its interior,  $y_1$  must lie on the left of  $R_1$  (as in Figure 6) or the top of  $R_1$ .

- Let  $R_4$  be the rectangle that is maximal with respect to the properties that it contains  $y_2$  at its bottom-right corner and  $y_1$  along its left edge.
- Let  $y_3$  be the singularity on the top edge of  $R_4$ .
- Let  $R_5$  be the rectangle that is maximal with respect to the properties that it contains  $y_3$  at its top-right corner and also contains  $y_1$  along its left edge.
- Let  $y_4$  be the singularity on the bottom edge of  $R_5$ .

In other words,  $R_4$  is obtained from the rectangle spanned by the saddle connection  $y_1 y_2$  and extending upward as far as possible and  $R_5$  is obtained from the rectangle spanned by  $y_1 y_3$  by extending it downward as far as possible.

To complete the proof we treat three cases. In each case we will choose a particular saddle connection  $\sigma$  of one of the  $\gamma_i$ . We will then find a saddle connection  $y_i y_j$  or a rectangle  $R_i$  with slope greater or equal to that of  $\sigma$  and with distance at most 5 from  $R$  in  $\mathcal{G}_{\widehat{X}}$ .

*Case 1:  $y_3$  lies to the left of  $\gamma_1$ .* For this case let  $\sigma$  be the saddle connection of  $\gamma_1$  that crosses  $\ell$ . There are two slightly different subcases, according to whether  $y_1$  lies on the left-hand edge or the top edge of  $R_1$ . We begin by treating the former case in detail and then explaining the minor changes needed for the latter case.

In both subcases, it follows from the definition of  $R_4$  that  $y_1$  and  $y_3$  lie on the left-hand and top edges of  $R_4$ , respectively. It then follows that  $y_1$  lies on the left-hand edge of  $R_5$ , and that  $y_3$  is the top-right corner of  $R_5$ . It must be that  $y_3$ , and hence the top edge of  $R_5$ , lies outside  $\Delta$ . In particular, it must be that  $\sigma$  intersects the right-hand edge of  $R_5$ . It also follows from the definitions that  $y_4$  lies on the bottom edge of  $R_5$ .

For the case where  $y_1$  lies on the left-hand edge of  $R_1$ , we claim that the following hold:

- $\sigma$  intersects the right-hand edge of  $R_5$
- $y_3$  is the top-right corner of  $R_5$ ,
- $y_4$  lies on the bottom edge of  $R_5$ ,
- $\sigma$  intersects the bottom edge of  $R_5$ ,

- along the bottom edge of  $R_5$ ,  $y_4$  lies to the right of  $\sigma$ .

We already verified the first three items. Since  $R_4$  contains no singularities in its interior and since  $R_1$  contains a singularity below  $\ell$ , it must be that the bottom of  $R_5$  passes through  $R_1$ , which gives the fourth item. As  $R_1$  contains no singularities in the interior, and has no singularities on its left edge besides  $y_1$ , the fifth and final item follows.

It follows from the claim that

$$\text{slope}(y_3y_4) \geq \min \text{slope}(\gamma_1).$$

The path

$$R, y_1y_2, R_4, y_1y_3, R_5, y_3y_4$$

in  $\mathcal{G}_{\hat{X}}$  has length 5, and so the lemma follows in the case where  $y_1$  lies on the left-hand edge of  $R_1$ .

For the case where  $y_1$  lies in the (interior of the) top edge of  $R_1$ , the proof is very similar. The changes are that the left-hand edge of  $R_5$  does not overlap the left-hand edge of  $R_4$ , and that the bottom edge of  $R_5$  can be lower than the bottom edge of  $R_1$  (this can happen when the singularity of  $R_1$  below  $\ell$  lies to the left of  $R_5$ ). The fourth and fifth items above become:

- $\sigma$  intersects the bottom edge or the left-hand edge of  $R_5$ ,
- if  $\sigma$  intersects the bottom edge of  $R_5$ ,  $y_4$  lies to the right of  $\sigma$  along this edge.

We can think of the case where  $\sigma$  crosses the left-hand edge of  $R_5$  as being a case where  $\sigma$  crosses the bottom edge of  $R_5$  very far to the left, then these two items are morally equivalent to their counterparts in the first subcase. As such, we still obtain that  $\text{slope}(y_3y_4) \geq \min \text{slope}(\gamma_1)$ , and the proof concludes as in the first subcase.

*Case 2:  $y_3$  lies to the right of  $\gamma_2$ ,  $y_2$  lies to the right of  $\gamma_2$ .* For this case let  $\sigma$  be the saddle connection of  $\gamma_2$  that crosses  $\ell$ . We claim that in this case  $y_2y_3$  is a saddle connection contained in  $R_2$ . Since  $R_2$  contains no singularities in its interior, it follows from the definition of  $y_2$  that  $y_2$  must lie along the right-hand edge of  $R_2$ . It then follows from the maximality condition on  $R_2$  and the fact that  $\Delta$  contains no singularities in its interior that  $R_2$  has a singularity along its top edge. From the definition of  $y_3$ , the fact that  $R_2$  contains no singularities in its interior, and the assumption that  $y_3$  lies to the right of  $\gamma_2$ , it must be that the singularity on the top edge of  $R_2$  is nothing other than  $y_3$ . The claim follows.

Given the claim, we have the following path in  $\mathcal{G}_{\hat{X}}$ :

$$R, y_1y_2, R_4, y_2y_3, R_2.$$

Since  $\text{slope}(R_2)$  is equal to the slope of the saddle connection of  $\gamma_2$  comprising the diagonal of  $R_2$ , the lemma follows in this case.

*Case 3:  $y_3$  lies to the right of  $\gamma_2$ ,  $y_2$  lies to the left of  $\gamma_2$ .* For this case let  $\sigma$  be the saddle connection of  $\gamma_2$  that crosses  $R_4$ . We claim that

$$\text{slope}(y_2y_3) \geq \text{slope}(\sigma).$$

To see this, we focus attention on the rectangle  $R_4$ . By construction,  $\gamma_2$  crosses  $R_4$  along its top and right-hand edges. Also,  $y_2$  is the bottom-right corner of  $R_4$  and  $y_3$  lies along the top edge of  $R_4$  to the right of  $\gamma_2$  (by the way Case 3 is defined). The claim follows.

We have the following path in  $\mathcal{G}_{\widehat{X}}$ :

$$R, y_1y_2, R_4, y_2y_3.$$

Applying the claim, the lemma follows in this case.  $\square$

**4.2. Small-slope geodesics and approximating rectangles.** The second statement of Lemma 4.1 is a consequence of the first part of the lemma and the following auxiliary lemma, which uses the tools of Section 3.

**Lemma 4.2.** *Let  $X$  be a singular Euclidean surface, let  $R$  be a maximal rectangle in  $\widehat{X}$ , let  $F = 2\text{flex}(X) - 1$ , let  $K = K(X)$  be the separatrix complexity, and let  $Z$  be a collection of pairwise non-crossing saddle connections in  $X$ . Assume that*

$$\min \text{slope}(Z) < F^{-(2K+1)} \text{slope}(R).$$

*Then there is a  $\sigma \in Z$  and a lift  $\tilde{\sigma} \subseteq \widehat{X}$  that crosses the vertical edges of  $R$ .*

*Proof.* It suffices to consider the case where  $Z$  is a singleton  $\{\sigma\}$ . Let  $\tilde{\sigma}_-$  be the saddle connection in  $R$  of smallest slope, that is, the one joining its vertical edges, and let  $\sigma_-$  be its image in  $X$ . Let  $R_-$  be the rectangle spanned by  $\tilde{\sigma}_-$ . The vertical edges of  $R_-$  are contained in those of  $R$ .

By Lemma 3.5 there is a saddle connection  $\rho$  in  $\mathbf{r}(\sigma)$  that has slope bounded above by  $\text{slope}(\sigma)$ . In particular,  $\text{slope}(\rho) \leq F^{-(2K+1)} \text{slope}(R)$ . By Lemma 3.3,  $\text{slope}(\rho) \leq F^{-2K} \text{slope}(\sigma_-)$ . Then by Lemma 3.4,  $\rho$  crosses  $\sigma_-$ .

Let  $\tilde{\rho}$  be a lift of  $\rho$  to  $\widehat{X}$  that crosses  $\tilde{\sigma}_-$ . Since (as in the previous paragraph)  $\text{slope}(\rho) \leq \text{slope}(\sigma_-)$ , since  $\tilde{\rho}$  spans a rectangle, it must be that  $\tilde{\rho}$  crosses the vertical edges of  $R_-$ .

Choose a lift  $\tilde{\sigma}$  of  $\sigma$  with  $\tilde{\rho} \in \mathbf{r}(\tilde{\sigma})$ . By the definition of  $\mathbf{r}(\tilde{\sigma})$ , there is a rectangle  $R_{\tilde{\rho}}$  in  $\widehat{X}$  that contains  $\tilde{\rho}$  and has a segment of  $\tilde{\sigma}$  as its diagonal.

Since  $R_{\tilde{\rho}}$  contains no singularities in its interior and  $\tilde{\rho}$  crosses the vertical edges of  $R_-$ , the diagonal of  $R_{\tilde{\rho}}$  must also cross the vertical edges of  $R_-$ . But this diagonal is contained in  $\tilde{\sigma}$  and so  $\tilde{\sigma}$  crosses the vertical edges of  $R_-$  and hence of  $R$ . This completes the proof.  $\square$

*Proof of Lemma 4.1(2).* By the assumptions in the statement and the definition of a  $k$ -approximating rectangle we have

$$\text{slope}(Z) < F^{-(2K+6)} \min\{\text{slope}(\gamma_1) \cup \text{slope}(\gamma_2)\} \leq F^{-(2K+1)} \text{slope}(R).$$

An application of Lemma 4.2 completes the proof.  $\square$

## 5. THE FORCING LEMMA

The goal of the section is to prove Proposition 5.1, which we call the forcing lemma. Roughly, it tells us that if we have a pseudo-Anosov mapping class  $f$  (hence a notion of slope) and a train track  $\tau$ , then  $\tau$  carries the curves of very small slope if and only if it carries the horizontal foliation for  $f$ .

For the statement, we recall some basic definitions regarding train tracks. Let  $\tau$  be a train track in a surface  $S$ . We say that  $\tau$  is filling if each of its complementary regions in  $S$  is a disk with at most one puncture; each such region can be regarded as a polygon. An extension of  $\tau$  is a train track  $\tau'$  with  $\tau$  a subtrack. If  $\tau'$  is obtained from  $\tau$  by adjoining diagonals of complementary polygons, then we say that  $\tau'$  is a diagonal extension. Next, we say that  $\tau$  is transversely recurrent if for every branch

$b$  of  $\tau$  there is a curve  $c$  in  $S$  so that  $c$  is transverse to  $b$  and forms no bigons with  $\tau$ . As in the introduction, if  $c$  is a curve (or foliation) carried by  $\tau$  then  $\tau[c]$  is the sub-train track whose branches are traversed by  $c$ . We say that  $c$  is fully carried by  $\tau$  if  $\tau[c] = \tau$ .

**Proposition 5.1.** *Let  $f \in \text{Mod}(S)$  be pseudo-Anosov with horizontal foliation  $\mathcal{F}_h$ , and let  $\tau$  be a filling, transversely recurrent train track in  $S$ . If  $c$  is a curve in  $S$  or a foliation in  $S$  with*

$$\mathcal{H}_f(c) \geq \mathcal{H}_f(\tau) + Q_{\text{force}},$$

*then the following hold.*

- (1) *If  $c < \tau$ , then  $\mathcal{F}_h < \tau[c]$ .*
- (2) *If  $\mathcal{F}_h$  is fully carried by  $\tau$ , then  $c$  is fully carried by a diagonal extension of  $\tau$ .*

We begin in Section 5.1 by defining left- and right-turning paths and explaining some of their properties. We then proceed to define the slope of a train track and relate it to the slopes of the left- and right-turning paths (Lemma 5.3). In Section 5.2 we introduce the notion of a shadow for the lift of a train track to the universal cover of the ambient surface. We then use the notion of a shadow to define obstructing branches and agreeable branches, and use these to give criterion for a foliation to be carried on a train track and a criterion for a carried foliation to traverse a branch of a train track, respectively. In Section 5.3 we prove Proposition 5.1 by first proving a version for train track branches, Lemma 5.6. The latter is an application of the approximating rectangles lemma, Lemma 4.1

**5.1. Left- and right-turning paths and slopes of train tracks.** The goal of this section is to define left- and right-turning paths for a lifted train track, and to compare their slopes to those of the corresponding train track. We begin with preliminaries from the theory of train tracks: transverse recurrence and quasi-geodesic train tracks.

*Transverse recurrence and quasi-geodesic train tracks.* If  $\tau$  is a transversely recurrent train track in a hyperbolic surface  $S$ , then the preimage  $\tilde{\tau}$  of  $\tau$  in  $\mathbb{H}^2$  is quasi-geodesic, in the sense that every bi-infinite train path in  $\tilde{\tau}$  is a quasi-geodesic with uniform quasi-isometry constants that do not depend on the path [24]. This fact is the main reason why we frequently employ the assumption on transverse recurrence.

Because of this, we may apply the tools and techniques from hyperbolic geometry to  $\tilde{\tau}$ . For instance, every bi-infinite train path  $\gamma$  in  $\tilde{\tau}$  has well-defined endpoints in  $\partial\mathbb{H}^2$ . For a singular Euclidean surface  $X$  there is a geodesic in the completed universal cover  $\hat{X}$  with the same endpoints as  $\gamma$  (this geodesic is unique except when  $\gamma$  corresponds to the lift of the core of a flat annulus). We will use the fact that two such paths  $\gamma_1$  and  $\gamma_2$  are linked at infinity if and only if their geodesic representatives in  $\hat{X}$  are.

As in the work of Masur–Minsky [19, Lemma 4.5], a transversely recurrent track  $\tilde{\tau} \subset \mathbb{H}^2$  does not admit any generalized bigons, by which we mean a region in  $\mathbb{H}^2$  bounded by either (i) distinct, bi-infinite train paths with common endpoints in  $\partial\mathbb{H}^2$  or (ii) distinct rays with common starting points in  $\mathbb{H}^2$  and common ending points in  $\partial\mathbb{H}^2$ .

*Left- and right-turning paths and cycles.* Let  $\tau$  be a train track in either a hyperbolic surface  $S$  or in the universal cover  $\mathbb{H}^2$ . Let  $b$  be a branch of  $\tau$ . The *left-turning path* in  $\tau$  through  $b$  is the bi-infinite train path  $\gamma_b^L$  in  $\tau$  that contains  $b$  and has the following property: if we travel along  $\gamma_b^L$  from  $b$  to another branch, then at each switch we traverse the left-most branch in the direction of travel.

When  $\tau$  is a train track in a surface  $S$  and  $b$  is an oriented branch in  $\tau$ , we may also define the *left-turning cycle*  $\beta_b^L$  via the following construction. We start from  $b$  and proceed in the direction of the orientation, always turning left, until we arrive at a branch  $b'$  that we have already traversed in the direction we are about to traverse it again. Then  $\beta_b^L$  is the closed train path obtained by discarding the initial path from  $b$  to  $b'$ . Right-turning paths and cycles  $\gamma_b^R$  and  $\beta_b^R$  are defined similarly.

We have the following basic facts about a transversely recurrent train track  $\tau$  in a surface  $S$ :

- for branch  $b$  of  $\tau$  the paths  $\gamma_b^L$  and  $\gamma_b^R$  are simple,
- for each oriented branch  $b$  of  $\tau$  the curves  $\beta_b^L$  and  $\beta_b^R$  are simple, and
- for any oriented edge  $b$ , the paths and curves  $\gamma_b^L$  and  $\beta_b^L$  are disjoint, and similarly for  $\gamma_b^R$  and  $\beta_b^R$ .

All three statements are verified by checking that the lifts to  $\mathbb{H}^2$  can be realized disjointly, equivalently that there are no two (components of) lifts whose endpoints separate each other in  $\partial\mathbb{H}^2$ . If this were the case, then at the first and last intersections of the first path with the second, the first path would be turning in the same direction with respect to the second, which is impossible.

One other fact that we will use is that each curve of the form  $\beta_b^L$  or  $\beta_b^R$  is a vertex cycle for  $\tau$ . This is straightforward from the definition; see, e.g. [14, Lemma 2.12].

*Slopes of train tracks versus slopes of left- and right-turning paths.* We define the slope of a train track  $\tau$  to be the union of the slopes of its vertex cycles, and we define  $\mathcal{H}_f(\tau)$  in the usual way from the slope. Recall that  $Q_{\text{curve}} = 2K^2$ , that  $E = 14$ , and that  $Q_{\text{track}} = EQ_{\text{curve}}$ . We prove now two lemmas in turn.

**Lemma 5.2.** *Let  $\tau$  be a transversely recurrent train track in a surface  $S$ . Then*

$$\text{diam } \mathcal{H}_f(\tau) \leq Q_{\text{track}}.$$

*Proof.* As in the introduction,  $E(S)$  is an upper bound for the diameter in  $\mathcal{C}(S)$  of the set of vertex cycles of  $\tau$ . By Proposition 3.1, each edge  $(v, w)$  in  $\mathcal{C}(S)$  has  $\text{diam } \mathcal{H}_f(v) \cup \mathcal{H}_f(w) \leq Q_{\text{curve}}$ . Finally, for any edge in the curve graph, the corresponding saddle connections are pairwise non-crossing (this is even true in the exceptional cases of  $S_{1,1}$  and  $S_{0,4}$  where edges do not correspond to disjoint curves but to curves that intersect once and twice, respectively). The lemma follows.  $\square$

**Lemma 5.3.** *Let  $f \in \text{Mod}(S)$  be pseudo-Anosov with singular Euclidean surface  $X$  and let  $K = K(X)$ . If  $\tau \subseteq X$  is a transversely recurrent train track then for each branch  $b$  of  $\tau$  we have*

$$\mathcal{H}_f(\tau) - (Q_{\text{track}} + Q_{\text{curve}}) \leq \mathcal{H}_f(\gamma_b^L) \leq \mathcal{H}_f(\tau) + (Q_{\text{track}} + Q_{\text{curve}})$$

and similarly for  $\gamma_b^R$ .

*Proof.* As above, we have that every  $\gamma_b^L$  is disjoint from the corresponding vertex cycle  $\beta_b^L$ . Again by Proposition 3.1 we have  $\text{diam } \mathcal{H}_f(\gamma_b^L) \cup \mathcal{H}_f(\beta_b^L) \leq Q_{\text{curve}}$ . By Lemma 5.2 and the fact that  $\beta_b^L$  is a vertex cycle of  $\tau$ , the lemma follows.  $\square$

**5.2. Obstructing and agreeable branches.** In this section we introduce the notion of a shadow for a train track. We then define the notion of an obstructing branch and use it to give a criterion for a foliation to be carried by a train track. Finally, we define the notion of an agreeable branch and use it to give a criterion for a carried foliation to traverse a certain branch of a train track.

*Shadows, obstructing branches, and agreeable branches.* Let  $\tau$  be a transversely recurrent train track in a hyperbolic surface  $S$  and  $\tilde{\tau}$  the preimage in  $\mathbb{H}^2$ . Let  $b$  be a branch of  $\tilde{\tau}$ . We define the shadow  $\Lambda_b$  as the union of the two disjoint intervals of  $\partial\mathbb{H}^2$  that run counterclockwise from an endpoint of  $\gamma_b^R$  to an endpoint of  $\gamma_b^L$ .

Next, let  $\gamma$  be a bi-infinite path in  $\mathbb{H}^2$  with well-defined endpoints at infinity (for instance a component of the preimage of a curve in  $S$  or a leaf of a foliation of  $S$ ). We say that  $b$  obstructs  $\gamma$  if the endpoints of  $\gamma$  in  $\partial\mathbb{H}^2$  separate the two components of the shadow  $\Lambda_b$  (in the language of Masur–Minsky [19, Proof of Lemma 4.5],  $b$  consistently separates  $\gamma$ ). We then say that  $b$  obstructs a foliation  $\mathcal{F}$  of  $S$  if it is obstructing for some component of the preimage of some singularity-free leaf of  $\mathcal{F}$ .

Finally, we say that the branch  $b$  is agreeable to the path  $\gamma$  if the endpoints of  $\gamma$  in  $\partial\mathbb{H}^2$  lie in different components of  $\Lambda_b$ . As in the obstructing case, we say that  $b$  is agreeable to a foliation  $\mathcal{F}$  if it is agreeable to some component of the preimage of some singularity-free leaf of  $\mathcal{F}$ .

*Obstructing branches and a criterion for carrying.* The next lemma uses obstructing branches to give a criterion for a curve or a foliation to be carried by a train track.

**Lemma 5.4.** *Let  $\tau$  be a filling transversely recurrent train track in a hyperbolic surface  $S$ , and let  $\mathcal{F}$  be a measured foliation of  $S$ . Then  $\mathcal{F}$  is carried on a diagonal extension of  $\tau$  if and only if  $\tau$  has no obstructing branches for  $\mathcal{F}$ .*

*Proof.* A proof of the reverse implication can be found in the work of Masur–Minsky [19, Proof of Lemma 4.5]. The argument there is framed in the special case where  $\mathcal{F}$  is a simple closed curve (as opposed to a general foliation) but the same argument applies. (The hypotheses of the Masur–Minsky lemma include the assumption that  $\tau$  is bi-recurrent; however, this part of the proof only uses the transverse recurrence.)

It remains to prove the forward implication. Suppose that  $\mathcal{F}$  is carried by  $\tau$ . Let  $\ell$  be a singularity-free leaf of the preimage of  $\mathcal{F}$  in  $\mathbb{H}^2$ . Also let  $\tilde{\tau}$  be the preimage of  $\tau$  in  $\mathbb{H}^2$  and let  $b$  be a branch of  $\tilde{\tau}$ . We would like to show that  $b$  is not an obstructing branch for  $\ell$ .

Let  $\gamma$  be the bi-infinite train path in  $\tilde{\tau}$  whose endpoints in  $\partial\mathbb{H}^2$  are the same as those of  $\ell$ . Consider now the intersection  $\gamma \cap \gamma_b^L$ . If this intersection is empty, then the endpoints of  $\ell$  and  $\gamma_b^L$  do not link in  $\partial\mathbb{H}^2$  and it follows that  $b$  is not obstructing for  $\ell$ .

Suppose then that  $b'$  is a branch of  $\gamma \cap \gamma_b^L$ . There are two sub-paths of  $\gamma$  that start at  $b'$ . Choose a sub-path  $\gamma'$  that moves away from  $b$  (if  $b = b'$  then both choices move away from  $b$ ). Since  $b'$  lies on  $\gamma_b^L$ , the endpoint of  $\gamma'$  at infinity lies in the shadow  $\Lambda_b$ . Thus  $b$  is not obstructing for  $\ell$ , as desired.  $\square$

*Agreeable branches and a criterion for edge traversal.* The next lemma uses agreeable branches to give a criterion for a carried curve or foliation to traverse a particular edge of a train track.

**Lemma 5.5.** *Let  $\tau \subset S$  be a transversely recurrent train track in a hyperbolic surface, let  $\tilde{\tau}$  be the preimage in  $\mathbb{H}^2$ , and let  $\gamma$  be a bi-infinite train path in  $\tilde{\tau}$ . Suppose that the endpoints of  $\gamma$  in  $\partial\mathbb{H}^2$  lie in different components of  $\Lambda_b$  for some branch  $b$  of  $\tilde{\tau}$ . Then  $\gamma$  traverses  $b$ .*

*Proof.* Let  $\alpha$  be the intersection of  $\gamma_b^L$  and  $\gamma_b^R$ . It must be that  $\alpha$  is a connected train path in  $\tilde{\tau}$  (otherwise  $\tilde{\tau}$ , hence  $\tau$ , would bound an embedded bigon). So  $\gamma_b^L \cup \gamma_b^R$  divides  $\mathbb{H}^2$  into four regions, two that meet  $\Lambda_b$  and two others. If  $\gamma$  does not traverse  $\alpha$ , then it must enter the interior of one of the latter two regions. The intersection of  $\gamma$  with this region is a train path that is either

- a compact train path,
- an infinite train path asymptotic to  $\partial\Lambda_b$ , or
- a bi-infinite train path (namely, all of  $\gamma$ ) asymptotic to  $\partial\Lambda_b$  at both ends.

In the first case, we see that  $\tilde{\tau}$ , hence  $\tau$ , bounds a bigon, and in the second and third cases we see that  $\tilde{\tau}$  has a generalized bigon. These contradictions complete the proof.  $\square$

**5.3. Proof of the forcing lemma.** In this section we prove Proposition 5.1, the forcing lemma. This proposition concludes that certain curves or leaves of a foliation are carried by a train track and also that every branch of the train track is traversed by this carrying. Before proceeding to the proof of the forcing lemma, we give a local version, that is, a version for a single branch of a (lifted) train track.

**Lemma 5.6.** *Let  $f \in \text{Mod}(S)$  be pseudo-Anosov with singular Euclidean surface  $X$  and horizontal foliation  $\mathcal{F}_h$ . Let  $\tau \subset S$  be a filling transversely recurrent train track and let  $\eta$  be a curve or a leaf of a foliation in  $S$  with*

$$\mathcal{H}_f(\eta) \geq \mathcal{H}_f(\tau) + Q_{\text{force}}.$$

*Let  $\tilde{\tau}$ ,  $\tilde{\mathcal{F}}_h$ , and  $\tilde{\eta}$  be the preimages in  $\mathbb{H}^2$  of  $\tau$ ,  $\mathcal{F}_h$ , and  $\eta$ . Let  $b$  be a branch of  $\tilde{\tau}$ , and let  $\ell$  be a singularity-free leaf of  $\tilde{\mathcal{F}}_h$ .*

- (1) *If  $b$  is agreeable to  $\ell$  then  $b$  is agreeable to some component of  $\tilde{\eta}$ .*
- (2) *If  $b$  is obstructing for  $\ell$  then  $b$  is obstructing for some component of  $\tilde{\eta}$ .*

*Proof.* We would like to apply Lemma 4.1 to the triple  $(\gamma_b^L, \gamma_b^R, \ell)$ , where the paths  $\gamma_b^L$  and  $\gamma_b^R$  have been replaced with their geodesic representatives in  $\hat{X}$ . To this end, we verify the hypotheses of that lemma. We have assumed in the hypotheses of this lemma that  $\text{flex}(X) < \infty$  and that  $\ell$  contains no singularities. We also have that  $\gamma_b^L$  and  $\gamma_b^R$  intersect, since they are the geodesic representatives of left- and right-turning paths through  $b$  (and hence must be linked at infinity).

We claim that  $\ell$  crosses both  $\gamma_b^L$  and  $\gamma_b^R$ . Indeed, the only other possibility is that an endpoint of  $\ell$  meets  $\Lambda_b$  in its boundary. This is impossible because the points of  $\partial\Lambda_b$  are endpoints of lifts of the vertex cycles of  $\tau$  corresponding to the left- and right-turning cycles through the image of  $b$ .

Since the hypotheses of Lemma 4.1 are satisfied, it follows from the first statement of the lemma that there is an  $F^5$ -approximating rectangle  $R$  for the triple  $(\gamma_b^L, \gamma_b^R, \ell)$  where  $F$  denotes the quantity  $2 \text{flex}(X) - 1$ .

We now would like to apply the second statement of Lemma 4.1 to  $\eta$ , which requires us to prove

$$\min \text{slope}(\eta) < F^{-(2K+6)} \min \{ \text{slope}(\gamma_b^L) \cup \text{slope}(\gamma_b^R) \}$$



By our hypothesis on  $\eta$  and Lemma 5.3, we have

$$\begin{aligned} \mathcal{H}_f(\eta) &\geq \mathcal{H}_f(\tau) + Q_{\text{force}} \\ &= \mathcal{H}_f(\tau) + Q_{\text{curve}} + Q_{\text{track}} + (2K + 6) \cdot 2K \\ &\geq \max\{\mathcal{H}_f(\gamma_b^L) \cup \mathcal{H}_f(\gamma_b^R)\} + (2K + 6) \cdot 2K \end{aligned}$$

Converting horizontalities to slope we have

$$\text{slope}(\eta) \leq (\lambda^{2K})^{-(2K+6)} \min\{\text{slope}(\gamma_b^L) \cup \text{slope}(\gamma_b^R)\}$$

Hence to verify the hypothesis for  $\eta$  it suffices to show that  $F < \lambda^{2K}$ . But this follows from Lemma 3.6 and Penner's inequality  $2 \leq \lambda^K$  from Section 3:

$$F = 2 \text{flex}(X) - 1 \leq 2\lambda^K - 1 < 2\lambda^K \leq \lambda^{2K}.$$

Having verified the hypothesis on  $\eta$ , Lemma 4.1(2) gives a geodesic  $\tilde{\eta}$  that is the geodesic representative of a component of the preimage of  $\eta$  and that crosses the vertical edges of  $R$ . Since  $R$  is an approximating rectangle, the geodesics  $\gamma_b^L$  and  $\gamma_b^R$  cross the horizontal edges of  $R$ . It follows that  $\tilde{\eta}$  crosses both  $\gamma_b^L$  and  $\gamma_b^R$  transversely within  $R$ .

In a CAT(0) space, the intersection of two geodesics is connected, and so  $\tilde{\eta}$  intersects each of  $\gamma_b^L$  and  $\gamma_b^R$  in a single point. Also, two geodesics that cross in a CAT(0) space cannot have the same endpoints at infinity. Therefore, the endpoints of  $\tilde{\eta}$  must be contained in the same components of  $\partial\mathbb{H}^2 \setminus (\partial\gamma_b^L \cup \partial\gamma_b^R)$  as the endpoints of  $\ell$ . This implies both statements of the lemma.  $\square$

*Proof of Proposition 5.1.* As in the statement,  $\mathcal{F}_h$  is the horizontal foliation for a pseudo-Anosov  $f \in \text{Mod}(S)$ . Also,  $\tau$  is a filling transversely recurrent train track in  $S$ , and  $c$  is a singularity-free leaf of a foliation in  $S$  with

$$\mathcal{H}_f(c) \geq \mathcal{H}_f(\tau) + Q_{\text{force}}.$$

In what follows,  $X$  is a singular Euclidean structure associated to  $f$ , and  $\tilde{\tau}$  and  $\tilde{\mathcal{F}}_h$  are the preimages of  $\tau$  and  $\mathcal{F}_h$  in the universal cover  $\tilde{X}$ . We prove the two statements of the proposition in turn.

*Statement (1).* Assume that  $c < \tau$ . We first claim that  $\mathcal{F}_h$  is carried by a diagonal extension  $\tau'$  of  $\tau$ . Assume not. By Lemma 5.4 there is a leaf  $\ell$  of  $\tilde{\mathcal{F}}_h$  and a branch  $b$  of  $\tilde{\tau}$  that is obstructing for  $\ell$ . Then by Lemma 5.6,  $c$  is obstructed by  $b$ . By Lemma 5.4,  $c$  is not carried by  $\tau$ , the desired contradiction.

We now show that  $\mathcal{F}_h$  is carried by  $\tau'[c]$ . Since  $\tau'[c] = \tau[c]$ , this will complete the proof of the first statement. Let  $b$  be a branch of  $\tilde{\tau}'$  traversed by some singularity-free leaf  $\ell$  of  $\tilde{\mathcal{F}}_h$ . The branch  $b$  is agreeable to  $\ell$ . By Lemma 5.6, the branch  $b$  is agreeable to some component  $\tilde{c}$  of the preimage of  $c$ . By Lemma 5.5, the train path of  $\tilde{\tau}'$  corresponding to  $\tilde{c}$  traverses the branch  $b$  and the first statement follows.

*Statement (2).* As per the second statement, we assume that  $\mathcal{F}_h$  is fully carried by  $\tau$ . Let  $b$  be a branch of  $\tilde{\tau}$  traversed by (a leaf of)  $\tilde{\mathcal{F}}_h$ . As in the proof of Statement (1), Lemma 5.6 implies there is an component of the preimage of  $c$  that traverses  $b$ . Since the components of  $\tilde{c}$  do not cross, it follows that  $b$  is not obstructing for any component of the preimage of  $c$ . Since  $b$  was arbitrary, Lemma 5.4 implies that  $c$  is carried by a diagonal extension  $\tau'$  of  $\tau$ . Applying Lemma 5.6 to  $\tau'$ , we conclude that  $c$  is fully carried by  $\tau'[\mathcal{F}_h]$ . Since the latter is equal to  $\tau$ , the proof is complete.  $\square$

## 6. THE FITTING LEMMA

The goal of this section is to prove the following, which we call the fitting lemma.

**Proposition 6.1.** *Let  $f \in \text{Mod}(S)$  be pseudo-Anosov with horizontal foliation  $\mathcal{F}_h$ , and let  $M \in \mathbb{R}$ . Let  $T$  be the set of transversely recurrent train tracks in  $S$  that carry  $\mathcal{F}_h$  and have horizontality at most  $M$ . There is an  $f$ -invariant train track  $\nu$  so that  $\nu$  is carried by each element of  $T$  and*

$$\mathcal{H}_f(\nu) \leq M + Q_{\text{fit}}.$$

The  $\nu$  in the conclusion of Proposition 6.1 is fitted in the sense that it lies between  $\mathcal{F}_h$  and each  $\tau \in T$  (in the sense that  $\mathcal{F}_h < \nu < \tau$ ) and it has horizontality close to  $\mathcal{H}_f(T)$ .

The starting point for the proof of Proposition 6.1 is the work of Agol [1]. Given a pseudo-Anosov  $f$ , Agol gives a bi-infinite sequence of train tracks and multi-splittings

$$\cdots \rightarrow \nu_i \rightarrow \nu_{i+1} \rightarrow \cdots \rightarrow \nu_{i+k} \rightarrow \cdots$$

and a  $k > 0$  so that

$$f(\nu_i) = \nu_{i+k}$$

for all  $i$ ; where by a multi-splitting we mean a collection of simultaneous train track splits along one or more large branches. In particular, each  $\nu_i$  is  $f$ -invariant and the  $\mathcal{H}_f(\nu_i)$  tend to  $\pm\infty$  as  $i \rightarrow \pm\infty$ . Also, each  $\nu_i$  fully carries the unstable foliation for  $f$ , and the set of large branches split at each step is precisely the set of branches of largest weight with respect to the unstable foliation. The train track  $\nu$  in the conclusion of Proposition 6.1 will be chosen from among the  $\nu_i$  in Agol's sequence.

Agol proves that if  $\tau$  is a train track carrying the horizontal foliation for  $f$ , then  $\tau$  splits to some  $\nu_i$  in the sequence [1, Theorem 3.5]. The basic idea of the proof of Proposition 6.1 is to use this fact to find a  $\nu_i$  that all the  $\tau_i$  split to, and then to show that we can move backwards along the sequence of  $\nu_i$  until the slope condition is also satisfied.

We divide the proof into two steps. In Section 6.1 we prove Lemma 6.4, the backfolding lemma, which gives conditions under which the fold of a train track carried by a train track  $\tau$  is again carried by  $\tau$ ; this is the main technical lemma of the section. Then in Section 6.2 we focus specifically on the Agol train tracks described above, and use them to prove Proposition 6.1.

**6.1. The backfolding lemma.** Before proceeding to the statement and the proof of the backfolding lemma, we introduce the following criterion for when a single fold does not disturb the carrying relation. In the statement, we say that a shadow  $\Lambda_b$  is contained in a shadow  $\Lambda_a$  by a component-respecting inclusion if each component of  $\Lambda_a$  contains one component of  $\Lambda_b$ .

**Lemma 6.2.** *Let  $S$  be a hyperbolic surface and let  $\sigma$  and  $\tau$  be transversely recurrent train tracks in  $S$  with  $\sigma < \tau$ . Let  $\sigma \mapsto \sigma'$  be a single fold creating a new large branch  $b'$ . Let  $\tilde{\sigma}'$  and  $\tilde{\tau}$  be the preimages of  $\sigma'$  and  $\tau$  in  $\mathbb{H}^2$ , and let  $\tilde{b}' \subseteq \tilde{\sigma}'$  be a lift of  $b'$  to  $\mathbb{H}^2$ . Suppose there is a branch  $\tilde{a}$  of  $\tilde{\tau}$  with*

$$\Lambda_{\tilde{b}'} \subseteq \Lambda_{\tilde{a}}$$

*with a component-respecting inclusion. Then  $\sigma'$  is carried by  $\tau$ .*

*Proof.* Let  $e$ ,  $f$ , and  $g$  be the branches that are folded together in  $\sigma$  to produce the branch  $b'$  of  $\sigma'$ , where  $f$  is the small branch. There are two cases: we can either choose the carrying map  $\sigma \rightarrow \tau$  so that a small neighborhood of  $f$  in  $\sigma$  maps homeomorphically or we cannot (here we are referring only to the carrying map, not the map to a tie neighborhood). In the latter case  $\sigma < \tau$  factors through  $\sigma \mapsto \sigma'$ , contrary to assumption. So we may assume that a small neighborhood of  $f$  does map homeomorphically; say that the branch in the image is  $h$ .

One way to visualize the statements in the preceding paragraph is to draw a tie neighborhood  $N$  of  $\tau$  and to consider  $\sigma \subset N$  so that collapsing along the fibers give a carrying map  $\sigma < \tau$ . If some train track tie of  $N$  passes through all three branches  $e$ ,  $f$ , and  $g$ , the carrying map  $\sigma < \tau$  factors through  $\sigma \mapsto \sigma'$ . The only other possibility is that  $e$ ,  $f$ , and  $g$  are configured in  $N$  as in Figure 7.

Let  $\tilde{e}$ ,  $\tilde{f}$ , and  $\tilde{g}$  be preimages of  $e$ ,  $f$ , and  $g$  in  $\tilde{\sigma}$  that are folded to create  $b'$ . Because of the folding relation, we have

$$\Lambda_{\tilde{e}} \cup \Lambda_{\tilde{f}} \cup \Lambda_{\tilde{g}} \subseteq \Lambda_{b'}$$

by component-respecting inclusions. By assumption we have an edge  $\tilde{a}$  of  $\tilde{\tau}$  with  $\Lambda_{b'} \subseteq \Lambda_{\tilde{a}}$  and hence  $\Lambda_{\tilde{a}}$  contains  $\Lambda_{\tilde{e}}$ ,  $\Lambda_{\tilde{f}}$ , and  $\Lambda_{\tilde{g}}$ . Let  $\tilde{h}$  be the branch of  $\tilde{\tau}$  that is the homeomorphic image of  $\tilde{f}$ .

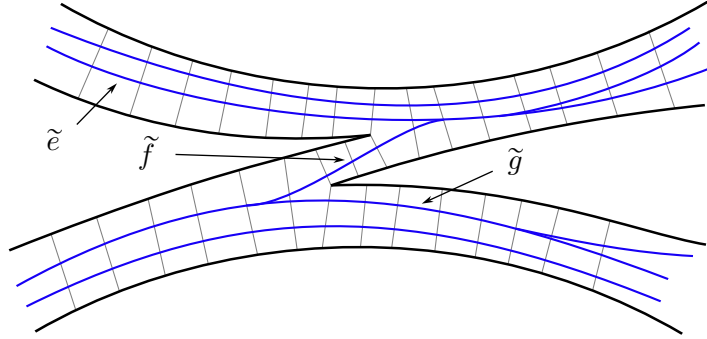


FIGURE 7. The configuration of  $\sigma \subset N$  if  $\sigma < \tau$  does not factor through  $\sigma \mapsto \sigma'$

By the above (component-respecting) inclusions, the path

$$\gamma = \gamma_{\tilde{f}}^L$$

has endpoints in both components of  $\Lambda_{\tilde{a}}$ . By Lemma 5.5, the image of  $\gamma$  in  $\tilde{\tau}$  passes through  $\tilde{a}$ . In particular,  $\tilde{h}$  and the branch  $\tilde{a}$  are connected by a train path in  $\tilde{\tau}$  (a sub-path of  $\gamma$ ).

Finally, observe from Figure 7 that  $\Lambda_{\tilde{h}}$  contains neither  $\Lambda_{\tilde{e}}$  nor  $\Lambda_{\tilde{g}}$ . Since moving along a train path does not increase the shadow in the forward direction, we must have that one of the inclusions  $\Lambda_{\tilde{e}} \subseteq \Lambda_{\tilde{a}}$  or  $\Lambda_{\tilde{g}} \subseteq \Lambda_{\tilde{a}}$  is false. This is the desired contradiction.  $\square$

We require one more lemma about slopes before proceeding to the backfolding lemma. Recall that  $E = 14$  and  $Q_{\text{curve}}(S) = 2K^2$ .

**Lemma 6.3.** *Let  $f \in \text{Mod}(S)$  be pseudo-Anosov, let  $\sigma \mapsto \sigma'$  be a single fold of transversely recurrent train tracks, and let  $b' \subseteq \sigma'$  be the large branch created in the fold. Then*

$$\min \mathcal{H}_f(\sigma) - (E + 3)Q_{\text{curve}} \leq \mathcal{H}_f(\gamma_{b'}^{L/R}) \leq \max \mathcal{H}_f(\sigma) + (E + 3)Q_{\text{curve}}.$$

*Proof.* By symmetry, it suffices to prove the first inequality. Let  $b$  be the branch of  $\sigma$  used to do the fold. Depending upon the orientation of the fold, it is either the case that  $\gamma_{b'}^L$  and  $\gamma_b^L$  are homotopic or  $\gamma_{b'}^R$  and  $\gamma_b^R$  are (up to modifying the fold itself by an ambient isotopy, we can in fact make the two paths equal). Say it is the former that are homotopic. As in the proof of Lemma 5.3, we have

$$\mathcal{H}_f(\gamma_{b'}^L) = \mathcal{H}_f(\gamma_b^L) \geq \min \mathcal{H}_f(\sigma) - Q_{\text{curve}}.$$

Using the same reasoning and applying Lemma 5.3, we have that

$$\mathcal{H}_f(\gamma_{b'}^R) \geq \mathcal{H}_f(\gamma_{b'}^L) - (E + 2)Q_{\text{curve}} \geq \min \mathcal{H}_f(\sigma) - (E + 2)Q_{\text{curve}}.$$

Combining these gives the lemma.  $\square$

We now arrive at the backfolding lemma.

**Lemma 6.4.** *Let  $f \in \text{Mod}(S)$  be pseudo-Anosov with singular Euclidean surface  $X$  and horizontal and vertical foliations  $\mathcal{F}_h$  and  $\mathcal{F}_v$ . Let  $\tau$  be a transversely recurrent train track in  $S$  that fully carries  $\mathcal{F}_h$ , and let  $\sigma$  be a train track in  $S$  with  $\sigma < \tau$  and*

$$\mathcal{H}_f(\sigma) \geq \mathcal{H}_f(\tau) + Q_{\text{fit}}$$

*Further suppose that  $\sigma \mapsto \sigma'$  is a single fold producing a large branch  $b' \subseteq \sigma'$  and that  $b'$  obstructs  $\mathcal{F}_v$ . Then  $\sigma'$  is carried by  $\tau$ .*

*Proof.* Our goal is apply Lemma 6.2. As such, our main work is to verify the hypotheses of that lemma, with  $\tau$ ,  $\sigma$ ,  $\sigma'$ , and  $b'$  in this lemma corresponding to the objects with the same names in the former. For starters, since  $f$  is pseudo-Anosov, we have that  $S$  is hyperbolic, as required by Lemma 6.2.

Let  $\tilde{b}'$  be a lift of  $b'$  to  $\mathbb{H}^2$ . The only hypothesis of Lemma 6.2 that it remains to verify is that there is a branch  $\tilde{a}$  in  $\tilde{\tau}$ , the preimage of  $\tau$  in  $\mathbb{H}^2$ , with

$$\Lambda_{\tilde{b}'} \subseteq \Lambda_{\tilde{a}}.$$

We first address the case where  $\tau$  is complete, meaning that  $\tau$  has no proper extensions, and then we use this special case to derive the general case.

*Complete case.* We first claim that  $\gamma_{b'}^L$  and  $\gamma_{b'}^R$  are carried by  $\tau$ . We will prove this using Proposition 5.1, and so we would like to bound the horizontalities of both paths. First, the assumption  $\mathcal{H}_f(\sigma) \geq \mathcal{H}_f(\tau) + Q_{\text{fit}}$  can be rephrased as  $\min \mathcal{H}_f(\sigma) \geq \mathcal{H}_f(\tau) + Q_{\text{fit}}$ . Recall that  $Q_{\text{fit}} = Q_{\text{force}} + (E + 3)Q_{\text{curve}}$ . Using Lemma 6.3 and the rephrasing of the assumption we have

$$\begin{aligned} \mathcal{H}_f(\gamma_{b'}^{L/R}) &\geq \min \mathcal{H}_f(\sigma) - (E + 3)Q_{\text{curve}} \\ &= \min \mathcal{H}_f(\sigma) - (Q_{\text{fit}} - Q_{\text{force}}) \\ &\geq \mathcal{H}_f(\tau) + Q_{\text{force}}. \end{aligned}$$

By Proposition 5.1, the assumption that  $\tau$  fully carries  $\mathcal{F}_h$  and the assumption that  $\tau$  is complete, the claim now follows.

By the previous claim the paths  $\gamma_{b'}^L$  and  $\gamma_{b'}^R$  in  $\mathbb{H}^2$  are carried by  $\tilde{\tau}$ . Also, these two paths are linked at infinity in  $\mathbb{H}^2$ . Thus, there is a branch  $\tilde{a}$  in  $\tilde{\tau}$  contained in both paths. We will show that this  $\tilde{a}$  satisfies the desired containment  $\Lambda_{\tilde{b}'} \subseteq \Lambda_{\tilde{a}}$ .

Since  $\gamma_{b'}^L$  and  $\gamma_{b'}^R$  both pass through  $\tilde{a}$ , we have

$$\partial\Lambda_{\tilde{b}'} \subseteq \Lambda_{\tilde{a}}.$$

From this containment, there are two possibilities for  $\Lambda_{\tilde{b}'}$ : either  $\Lambda_{\tilde{b}'} \subseteq \Lambda_{\tilde{a}}$  (which is what we want to show) or  $\partial\Lambda_{\tilde{a}} \subseteq \Lambda_{\tilde{b}'}$ . As such, it remains to rule out the latter.

By assumption there is a leaf  $\ell$  of  $\mathcal{F}_v$  that separates the components of  $\Lambda_{\tilde{b}'}$ . Any such  $\ell$  must cross both  $\gamma_{b'}^L$  and  $\gamma_{b'}^R$ . By (the unstated vertical version of) Lemma 4.1 there is a maximal rectangle  $R$  so that  $\ell$  meets  $R$  in a segment, so that the geodesic representative of  $\gamma_{b'}^L$  in  $\hat{X}$  crosses the vertical edges of  $R$ , and so that every geodesic  $\gamma$  in  $X$  with

$$\max \text{slope}(\gamma) > F^{2K+6} \text{slope}(\gamma_{b'}^i) \text{ for } i \in \{L, R\}$$

there is a component  $\tilde{\gamma}$  of the preimage of  $\gamma$  the crosses the horizontal edges of  $R$ . Here  $F = 2 \text{flex}(X) - 1$ , as usual.

We next claim that  $\gamma_{\tilde{a}}^L$  satisfies the criterion

$$\max \text{slope}(\gamma_{\tilde{a}}^L) > F^{2K+6} \text{slope}(\gamma_{b'}^i) \text{ for } i \in \{L, R\}.$$

First, by Proposition 3.1, the assumption on  $\mathcal{H}_f(\sigma)$ , Lemma 6.3, the definition of  $E$ , and the fact that  $Q_{\text{fit}} \geq 0$ , we have

$$\begin{aligned} \min \mathcal{H}_f(\gamma_{\tilde{a}}^L) &\leq \max \mathcal{H}_f(\tau) + Q_{\text{curve}} \\ &\leq \min \mathcal{H}_f(\sigma) + Q_{\text{curve}} - Q_{\text{fit}} \\ &\leq \mathcal{H}_f(\gamma_{b'}^{L/R}) + (E + 4)Q_{\text{curve}} - Q_{\text{fit}} \\ &= \mathcal{H}_f(\gamma_{b'}^{L/R}) + 18Q_{\text{curve}} - Q_{\text{fit}} \\ &\leq \mathcal{H}_f(\gamma_{b'}^{L/R}) + 18Q_{\text{curve}}. \end{aligned}$$

We convert to slopes and use Penner's inequality  $\lambda^{2K} > F$  from Section 5:

$$\begin{aligned} \max \text{slope}(\gamma_{\tilde{a}}^L) &\geq (\lambda^{2K})^{18K} \text{slope}(\gamma_{b'}^{L/R}) \\ &\geq F^{18K} \text{slope}(\gamma_{b'}^{L/R}). \end{aligned}$$

Since  $18K > 2K + 6$ , the claim follows. (This argument does not use  $\mathcal{H}_f(\sigma) \geq \mathcal{H}_f(\tau) + Q_{\text{fit}}$ , but the weaker assumption that  $\mathcal{H}_f(\sigma) \geq \mathcal{H}_f(\tau) + (E + 2)Q_{\text{curve}} - 12K$ .)

Let us replace  $\gamma_{\tilde{a}}^L$ ,  $\gamma_{b'}^L$ , and  $\gamma_{b'}$  with their geodesic representatives in  $\hat{X}$ . By the previous claim and the description of  $R$  from Lemma 4.1 (plus the CAT(0) geometry of  $\hat{X}$ ) there is a translate  $\gamma'$  of  $\gamma = \gamma_{\tilde{a}}^L$  that crosses both  $\gamma_{b'}^L$  and  $\gamma_{b'}^R$  each in exactly one point. It follows that  $\gamma'$  also separates the components of  $\Lambda_{\tilde{b}'}$ . However, under our assumption  $\gamma$  has its boundary in  $\Lambda_{\tilde{b}'}$  and so  $\gamma$  and  $\gamma'$  must cross. This contradicts the fact that left turning paths are simple and we are done.

*General case.* Suppose that  $\sigma' \rightarrow \sigma$  is a left split, so that  $\gamma_b^L$  is automatically carried by  $\tau$ . Now by Proposition 5.1 and the assumption that  $\tau$  fully carries  $\mathcal{F}_h$ , we have that  $\gamma_b^R$  is carried by  $\tilde{\tau}'$ , where  $\tau'$  is some diagonal extension of  $\tau$ . Let  $\tilde{a}$  be a common branch of the train paths  $\gamma_b^{L/R}$  as carried by  $\tilde{\tau}'$ . As  $\gamma_b^L < \tilde{\tau}$ , we must have that  $\tilde{a}$  is contained in  $\tilde{\tau}$ . Note that since a diagonal of a diagonal extension is never the left/right-most turn, we have that  $\Lambda_{\tilde{a}}$  is independent of whether  $\tilde{a}$  is considered as a branch of  $\tilde{\tau}$  or  $\tilde{\tau}'$ . By the special case of the lemma, we have  $\Lambda_{\tilde{b}} \subseteq \Lambda_{\tilde{a}}$  and we are done.  $\square$

**6.2. Proof of the fitting lemma.** Before proving Proposition 6.1 we have a technical lemma about Agol’s train tracks, Lemma 6.5. This lemma will use a fact from the work of Guéritaud, Minsky, and the third-named author, which we now explain.

*Guéritaud and transversality.* Using the work of Guéritaud, Minsky and the third author observed that each Agol train track  $\nu_i$  can be realized as a smoothing of the dual graph to a triangulation of the surface by rectangle-spanning saddle connections. Moreover, the large half-branch in each triangle meets the tallest side. See Figure 8 for an illustration of the local picture in one such triangle. It follows that  $\nu_i$  can be chosen to be transverse to the vertical foliation  $\mathcal{F}_v$  for the corresponding pseudo-Anosov mapping class.

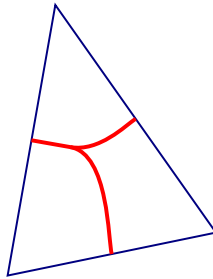


FIGURE 8. A piece of an Agol train track inside a triangle of a triangulation by rectangle-spanning saddle connections

**Lemma 6.5.** *Let  $f \in \text{Mod}(S)$  be pseudo-Anosov with vertical foliation  $\mathcal{F}_v$ , and let  $\nu$  be an Agol train track for  $f$ . Then every branch of  $\nu$  obstructs  $\mathcal{F}_v$ .*

*Proof.* Fix a singular Euclidean structure on  $S$  associated to  $f$ , and let  $\tilde{\nu}$  be the preimage of  $\nu$  in the corresponding completed universal cover  $\hat{X}$ . Let  $\tilde{\mathcal{F}}_v$  be the induced vertical foliation of  $\hat{X}$ . As above, we may assume that  $\tilde{\nu}$  is transverse to  $\tilde{\mathcal{F}}_v$ .

We claim that a train path in  $\tilde{\nu}$  meets a nonsingular leaf of  $\tilde{\mathcal{F}}_v$  at most once. Indeed, suppose that  $\ell$  is a nonsingular leaf of  $\tilde{\mathcal{F}}_v$  that meets some train path of  $\tilde{\nu}$  in more than one point. There is an embedded (innermost) disk  $D$  in  $\mathbb{H}^2$  bounded by a train path  $p$  in  $\tilde{\nu}$  and a segment of  $\ell$ , and so that the interior of  $D$  does not meet  $\tilde{\nu}$ . The disk  $D$  is foliated by (the restriction of)  $\tilde{\mathcal{F}}_v$ . If we double  $D$  across the segment of  $\ell$ , we obtain foliated disk where the leaves are transverse to the boundary. Doubling this resulting disk along its boundary, we obtain a foliated sphere where each singularity corresponds to a singularity in the interior of  $D$ . Each

such singularity has negative index, contradicting the Gauss–Bonnet formula. The claim follows.

Now, let  $b$  be a branch of  $\tilde{\nu}$ , and let  $\ell$  be a nonsingular leaf of  $\mathcal{F}_v$  that passes through  $b$ . By the claim,  $\ell$  meets  $\gamma_b^L$  and  $\gamma_b^R$  exactly once each. Thus the endpoints of  $\ell$  separate the components of  $\Lambda_b$  and so  $b$  obstructs  $\ell$ , hence  $\mathcal{F}_v$ , as required.  $\square$

*Proof of Proposition 6.1.* As above, the  $\mathcal{H}_f(\nu_i)$  tend to  $\pm\infty$  as  $i \rightarrow \pm\infty$ . Thus there is an  $i \in \mathbb{Z}$  so that  $\mathcal{H}_f(\nu_i) \leq M + Q_{\text{fit}}$  and  $\mathcal{H}_f(\nu_{i+k}) \geq M + Q_{\text{fit}}$  for all  $k > 0$ . To prove the proposition, it suffices to show that  $\nu_i$  is carried by all elements of  $T$ .

So let  $\tau \in T$ . As above, this means that  $\tau$  is a transversely recurrent train track in  $S$  that carries  $\mathcal{F}_h$  and has  $\mathcal{H}_f(\tau) \leq M$ . Let  $\tau'$  denote  $\tau[\mathcal{F}_h]$ , the sub-train track of  $\tau$  fully carrying  $\mathcal{F}_h$ . We will show that  $\nu_i < \tau'$ .

By the work of Agol discussed at the start of the section,  $\tau'$  splits to some  $\nu_j$  in the Agol sequence. Assume that  $j$  is the smallest integer with  $\nu_j < \tau'$ . If  $j \leq i$  we are done, because then  $\nu_i < \nu_j < \tau'$ . So we assume  $j > i$ . By our assumption in the first paragraph of the proof,  $\mathcal{H}_f(\nu_j) \geq M + Q_{\text{fit}}$ . Lemma 6.5 implies that every branch of  $\nu_j$  obstructs the vertical foliation  $\mathcal{F}_v$  for  $f$ .

Lemma 6.4 then implies that each individual fold of the multifold  $\nu_j \mapsto \nu_{j-1}$  produces a train track carried by  $\tau'$ . Since these individual folds are far apart from each other in the sense that the large branches they create are disjoint, this further implies that  $\nu_{j-1}$  is carried by  $\tau'$ , contradicting the minimality of  $j$ . Thus,  $j \leq i$  and so  $\nu_i < \nu_j < \tau'$ , as desired.  $\square$

## 7. PROOF OF THEOREM 1.1

Before proving our main theorem, we require four technical lemmas, a criterion for carrying, an upper bound on the radius of a Dehn–Thurston cell structure, an upper bound for the radius of the set of vertex cycles in a Dehn–Thurston train track, and a linear algebraic lemma about projective sinks.

*Another carrying criterion.* The following lemma gives a criterion for the extension operation to preserve the carrying relation between train tracks.

**Lemma 7.1.** *Suppose that  $\nu < \tau$ , where  $\nu$  and  $\tau$  are large train tracks in a surface. Let  $\tilde{\nu}$  be a diagonal extension of  $\nu$  that fully carries a curve  $d$  that is also carried by  $\tau$ . Then a carrying map  $\nu \rightarrow \tau$  extends to a carrying map  $\tilde{\nu} \rightarrow \tau$ .*

*Proof.* A carrying map from  $\nu$  to  $\tau$  takes  $\tilde{\nu}$  to a diagonal extension  $\tilde{\tau}$  of  $\tau$  (which we assume to be a minimal such extension). Since  $d$  is carried by  $\tau$  and  $\tilde{\nu}$  and since the train path induced by  $d$  is unique, it follows that  $\tilde{\tau} = \tau$ , as desired.  $\square$

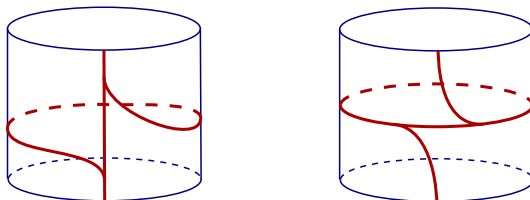


FIGURE 9. Two model train tracks on an annulus carrying a common path connecting the boundaries



*The radius for a Dehn–Thurston cell structure.* The following lemma justifies our choice of  $D(S) = 1$  in the introduction. Again, see the book by Penner–Harer for background on Dehn–Thurston train tracks [24, §2.6].

**Lemma 7.2.** *Let  $S$  be a surface, let  $Y$  be a Dehn–Thurston cell structure on  $\text{PMF}(S)$ , and let  $c$  be a curve of the underlying pants decomposition. Then the radius of  $Y$  with respect to  $c$  is 1.*

*Proof.* There are two types of train tracks in the given collection of Dehn–Thurston train tracks: (1) those that contain  $c_1$  as a vertex cycle and (2) those that do not carry  $c_1$ . To prove the lemma we must show that each train track of the second type carries a curve in common with some train track of the first type.

So let  $\tau$  be of the second type. This means that the model train track  $\tau_A$  for  $\tau$  in the annulus  $A$  around  $c_1$  does not carry  $c_1$ . Of the two model tracks on  $A$  that carry  $c_1$ , there is exactly one with the property that there is a homotopy class  $\alpha$  of paths in  $A$  (relative to  $\partial A$ ) carried both by this model and  $\tau_A$ ; see Figure 9. Let  $\tau'$  be the train track obtained from  $\tau$  by changing the model on  $A$  to this model.

Let  $c$  be a curve in  $S$  that is carried by  $\tau$  and so that each component of the intersection with  $A$  is parallel to  $\alpha$  (relative to  $\partial A$ ); for instance, there is a vertex cycle for  $\tau$  that passes through  $A$  once and satisfies this property. Then  $c$  is also carried by  $\tau'$  since  $\tau$  agrees with  $\tau'$  outside  $A$ . This completes the proof.  $\square$

*Sinks and eigenvalues.* Finally, we give a lemma that allows us to conclude the second statement of Theorem 1.1 from the first. As in the introduction, a vertex ray for a integral cone in an integral cone complex  $Y$  is a vertex of the corresponding polyhedron in  $\mathcal{P}(Y)$ .

**Lemma 7.3.** *Suppose that  $Y$  is an integral cone complex and that the group  $G$  acts on  $Y$  with property  $Q$ . Let  $f : Y \rightarrow Y$  be a projectively source-sink integral cone complex map with sink  $v$ , associated eigenvalue  $\lambda$ , and extended dynamical map given by  $(V_i, D)$ . Suppose that  $v$  does not lie on a rational ray in  $V_i$ . Then the largest real eigenvalue of  $D$  is equal to  $\lambda$  and any corresponding positive eigenvector represents  $v$ .*

*Proof of Lemma 7.3.* As in the extended dynamical matrix algorithm, the extended dynamical matrix  $D$  induces a linear map of a subspace  $W$  of  $\mathbb{R}^d$ , and moreover  $v$  lies in  $W$ . Suppose that  $w$  is an eigenvector of  $D$  with real eigenvalue  $\mu > 1$ . Then  $D$  acts on the subspace  $W'$  of  $W$  spanned by  $v$  and  $w$ . Also, by the description of  $D$  from the extended dynamical matrix algorithm,  $w$  lies in the subspace of  $\mathbb{R}^d$  spanned by  $V_i$ .

Let  $U \subseteq V_i$  be the PL-eigenregion used to determine  $D$ . Since  $v$  is not rational, it does not lie in a vertex ray of  $U$ , the intersection  $W' \cap U$  contains a two-dimensional cone  $C$  (for the same reason that the intersection of a line with a polytope is either a vertex or contains an open interval).

The actions of  $f$  and  $D$  agree on  $C$ . Since  $f$  acts with projective source-sink dynamics on  $Y$ , we can choose a smaller cone  $C'$  around the span of  $v$  with the property that  $f(C')$  is properly contained in  $C'$ . It follows that  $\mu < \lambda$ , as desired. The lemma follows.  $\square$

*Proof of Theorem 1.1.* We begin with the first statement. Let  $\{\tau_i\}$  be a set of complete, transversely recurrent train tracks giving the integral cone complex structure

$Y$  on  $\text{MF}(S)$ , and say that  $c$  is a vertex cycle of  $\tau_0$ . As per the statement of the theorem, we let  $C$  be a linear cone for the  $f$ -action on  $Y$  that contains  $f^Q(c)$ . We will show that  $C$  shares an open subset with a Nielsen–Thurston eigenregion for  $f$ .

Suppose  $C \subseteq V(\tau_i)$  and that  $f$  maps  $C$  linearly into  $V(\tau_j)$  (while our proof applies in all cases, the proof simplifies in the case  $i = j$ , which is the generic case; we leave this simplification to the reader). We claim that both  $V(\tau_i)$  and  $V(\tau_j)$  contain  $\mathcal{F}_h$ . Using the basic properties of horizontality, the definition of  $Q$ , and the definition of  $Q_{\text{DT}}$  with Proposition 3.1, we have:

$$\begin{aligned} \mathcal{H}_f(f^Q(c)) &= \mathcal{H}_f(c) + Q \\ &= \mathcal{H}_f(c) + Q_{\text{force}} + Q_{\text{fit}} + Q_{\text{DT}} + 1 \\ &\geq \mathcal{H}_f(\tau_i) + Q_{\text{force}} + Q_{\text{fit}} + 1, \end{aligned}$$

and similarly for  $\tau_j$ . As  $Q_{\text{force}} + Q_{\text{fit}} + 1 > Q_{\text{force}}$ , the claim now follows from Proposition 5.1.

Let

$$M = \max \mathcal{H}_f(\tau_i) \cup \mathcal{H}_f(\tau_j).$$

By the previous claim, Proposition 6.1 now gives an  $f$ -invariant train track  $\nu$  that fully carries  $\mathcal{F}_h$ , that is carried by both  $\tau_i$  and  $\tau_j$ , and that satisfies

$$\mathcal{H}_f(\nu) \leq M + Q_{\text{fit}}.$$

We now claim that

$$\mathcal{H}_f(f^Q(c)) > \mathcal{H}_f(\nu) + Q_{\text{force}}.$$

Indeed, let  $\tau$  be the element of  $\{\tau_i, \tau_j\}$  realizing  $M$ . Using the inequality obtained in the last claim, the definition of  $M$ , and the definition of  $\nu$  we have

$$\begin{aligned} \mathcal{H}_f(f^Q(c)) &\geq \max \mathcal{H}_f(\tau) + Q_{\text{force}} + Q_{\text{fit}} + 1 \\ &= M + Q_{\text{force}} + Q_{\text{fit}} + 1 \\ &\geq \mathcal{H}_f(\nu) + Q_{\text{force}} + 1, \end{aligned}$$

and the claim follows.

As horizontality is a continuous function on  $\text{MF}(S)$ , there is a neighborhood of  $f^Q(c)$  in  $V(\tau_i)$  consisting of points with horizontality greater than  $\mathcal{H}_f(\nu) + Q_{\text{force}}$ . Since  $\nu$  fully carries  $\mathcal{F}_h$ , it follows from Proposition 5.1 that each point of this neighborhood is carried by a diagonal extension of  $\nu$  (there is a simplification in the case where  $\nu$  is complete, which is the generic case; we again leave the simplification to the reader).

What is more, there is a neighborhood  $U$  of  $f^Q(c)$  in  $V(\tau_i)$  so that each point of  $U$  is carried by a diagonal extension of  $\nu$  that also carries  $f^Q(c)$ . This is because there are finitely many diagonal extensions of  $\nu$  and if one of these does not carry  $f^Q(c)$  the corresponding cone is not needed to carry a small neighborhood of  $f^Q(c)$ .

Let then  $d$  be a curve that lies in the interior of  $U$  and let  $\tilde{\nu}$  be a diagonal extension of  $\nu$  that fully carries  $d$ . Since the interior of  $U$  is (of course) open, we can further choose  $d$  so that  $\tilde{\nu}$  is complete and (by the previous paragraph) and also carries  $f^Q(c)$  (the set of curves carried by a non-complete train track has positive codimension). In particular,  $\tilde{\nu}$  is the unique diagonal extension of  $\nu$  that contains  $d$ .

The cones  $V(\tilde{\nu})$  and  $C$  have overlapping interiors, since they both contain a small neighborhood of  $d$  (using the fact that  $d$  is fully carried by  $\tilde{\nu}$ ).

We claim that  $f(\tilde{\nu}) < \tau_j$ . Indeed, first of all we have  $\nu < \tau_j$  so  $f(\nu) < \tau_j$ . Also, since  $f(C)$  lies in  $V(\tau_j)$  we have  $f(d) < \tau_j$ . Since  $\tilde{\nu}$  fully carries  $d$ ,  $f(\tilde{\nu})$  fully carries  $f(d)$ , and so Lemma 7.1 implies that  $\tau_j$  carries the unique diagonal extension of  $f(\nu)$  that fully carries  $f(d)$ . But this unique extension must be  $f(\tilde{\nu})$ , since  $\tilde{\nu}$  is the unique extension of  $\nu$  fully carrying  $d$  and since  $f$  is (represented by) a homeomorphism.

By the last claim that  $V(\tilde{\nu})$  a Nielsen–Thurston eigenregion for  $f$ . Indeed,  $V(\tilde{\nu})$  contains  $\mathcal{F}_h$  since  $V(\nu)$  does, and the previous claim implies that  $f$  is linear on  $V(\tilde{\nu})$ .

We have proved that the cone  $C$  and the Nielsen–Thurston eigenregion  $V(\tilde{\nu})$  are subcones of  $V(\tau_i)$  with overlapping interiors. Moreover,  $f$  maps each cone linearly into  $V(\tau_j)$  and so these linear actions must be equal. In particular, there is a single  $f$ -linear piece  $C_{NT}$  containing both. The region  $C_{NT}$  is itself a Nielsen–Thurston eigenregion, and the proof of the first statement is complete.

The second statement follows from the first statement and Lemma 7.3. Indeed, for the  $\text{Mod}(S)$ -action on the Dehn–Thurston integral cone complex, the foliation  $\mathcal{F}_h$  automatically satisfies the condition of not lying in a vertex ray. The reason is that in constructing the linear pieces for  $f$  as in the basic computational algorithm, the subdivisions of the cells of  $Y$  are made using rational planes. It follows that the vertices of the resulting subdivision of  $\mathcal{P}(Y)$  are rational, which means they correspond to multicurves in  $S$ . On the other hand, the unstable foliation for a pseudo-Anosov mapping class does not correspond to a multicurve. This completes the proof.  $\square$

## REFERENCES

- [1] Ian Agol. Ideal triangulations of pseudo-Anosov mapping tori. In *Topology and geometry in dimension three*, volume 560 of *Contemp. Math.*, pages 1–17. Amer. Math. Soc., Providence, RI, 2011.
- [2] Tarik Aougab. Uniform hyperbolicity of the graphs of curves. *Geom. Topol.*, 17(5):2855–2875, 2013.
- [3] Filippo Baroni. Uniformly polynomial-time classification of surface homeomorphisms, 2024.
- [4] Mark C. Bell and Saul Schleimer. Slow north-south dynamics on  $\mathcal{PML}$ . *Groups Geom. Dyn.*, 11(3):1103–1112, 2017.
- [5] Mark C. Bell and Richard C. H. Webb. Polynomial-time algorithms for the curve graph, 2016.
- [6] Mark Christopher Bell. *Recognising Mapping Classes*. PhD thesis, University of Warwick, 2015.
- [7] M. Bestvina and M. Handel. Train-tracks for surface homeomorphisms. *Topology*, 34(1):109–140, 1995.
- [8] Brian H. Bowditch. Uniform hyperbolicity of the curve graphs. *Pacific J. Math.*, 269(2):269–280, 2014.
- [9] Peter Brinkmann. An implementation of the Bestvina-Handel algorithm for surface homeomorphisms. *Experiment. Math.*, 9(2):235–240, 2000.
- [10] Andrew J. Casson and Steven A. Bleiler. Automorphisms of surfaces. Notes.
- [11] Benson Farb and Dan Margalit. *A primer on mapping class groups*, volume 49 of *Princeton Mathematical Series*. Princeton University Press, Princeton, NJ, 2012.
- [12] Albert Fathi, François Laudenbach, and Valentin Poénaru. *Travaux de Thurston sur les surfaces*, volume 66 of *Astérisque*. Société Mathématique de France, Paris, 1979. Séminaire Orsay, With an English summary.
- [13] Matthew D. Finn and Jean-Luc Thiffeault. Topological entropy of braids on the torus. *SIAM J. Appl. Dyn. Syst.*, 6(1):79–98 (electronic), 2007.
- [14] Vaibhav Gadre and Saul Schleimer. The curves not carried. *Groups Geom. Dyn.*, 10(4):1249–1264, 2016.
- [15] Vaibhav Gadre and Chia-Yen Tsai. Minimal pseudo-Anosov translation lengths on the complex of curves. *Geom. Topol.*, 15(3):1297–1312, 2011.
- [16] Toby Hall. Dynn (computer software). Version 0.10.3, 2015.

- [17] Toby Hall and S. Öykü Yurttaş. On the topological entropy of families of braids. *Topology Appl.*, 156(8):1554–1564, 2009.
- [18] William Jaco and Jeffrey L. Tollefson. Algorithms for the complete decomposition of a closed 3-manifold. *Illinois J. Math.*, 39(3):358–406, 1995.
- [19] Howard A. Masur and Yair N. Minsky. Geometry of the complex of curves. I. Hyperbolicity. *Invent. Math.*, 138(1):103–149, 1999.
- [20] Christof Menzel and John R. Parker. Pseudo-Anosov diffeomorphisms of the twice-pictured torus. In J. R. Cho and J. Mennicke, editors, *Recent Advances in Group Theory and Low-Dimensional Topology*, number 27, pages 141–154. Heldermann Verlag, research and exposition in mathematics edition, 2003. unpublished.
- [21] Yair N. Minsky and Samuel J. Taylor. Fibered faces, veering triangulations, and the arc complex. *Geom. Funct. Anal.*, 27(6):1450–1496, 2017.
- [22] Jacques-Olivier Moussaïf. On computing the entropy of braids. *Funct. Anal. Other Math.*, 1(1):37–46, 2006.
- [23] R. C. Penner. Probing mapping class groups using arcs. In *Problems on mapping class groups and related topics*, volume 74 of *Proc. Sympos. Pure Math.*, pages 97–114. Amer. Math. Soc., Providence, RI, 2006.
- [24] R. C. Penner and J. L. Harer. *Combinatorics of train tracks*, volume 125 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 1992.
- [25] Robert C. Penner. Bounds on least dilatations. *Proc. Amer. Math. Soc.*, 113(2):443–450, 1991.
- [26] Robert Clark Penner. *A COMPUTATION OF THE ACTION OF THE MAPPING CLASS GROUP ON ISOTOPY CLASSES OF CURVES AND ARCS IN SURFACES*. ProQuest LLC, Ann Arbor, MI, 1982. Thesis (Ph.D.)—Massachusetts Institute of Technology.
- [27] Kasra Rafi. A characterization of short curves of a Teichmüller geodesic. *Geom. Topol.*, 9:179–202, 2005.
- [28] Robert Tang and Richard C. H. Webb. Shadows of Teichmüller discs in the curve graph. *Int. Math. Res. Not. IMRN*, (11):3301–3341, 2018.
- [29] William P. Thurston. On the geometry and dynamics of diffeomorphisms of surfaces. *Bull. Amer. Math. Soc. (N.S.)*, 19(2):417–431, 1988.
- [30] Richard Webb. Personal communication, August 2024.
- [31] Chee-Keng Yap et al. *Fundamental problems of algorithmic algebra*, volume 49. Oxford University Press Oxford, 2000.